

Improved search for D^0 - \bar{D}^0 mixing using semileptonic decays at Belle

U. Bitenc,¹² I. Adachi,⁷ H. Aihara,⁴⁰ K. Arinstein,¹ T. Aushev,^{16,11} A. M. Bakich,³⁶ V. Balagura,¹¹ I. Bedny,¹ K. Belous,¹⁰ A. Bondar,¹ A. Bozek,²⁵ M. Bračko,^{18,12} J. Brodzicka,⁷ T. E. Browder,⁶ M.-C. Chang,³ A. Chen,²² W. T. Chen,²² B. G. Cheon,⁵ R. Chistov,¹¹ I.-S. Cho,⁴⁵ Y. Choi,³⁵ J. Dalseno,¹⁹ M. Danilov,¹¹ M. Dash,⁴⁴ A. Drutskoy,² S. Eidelman,¹ S. Fratina,¹² B. Golob,^{17,12} H. Ha,¹⁴ J. Haba,⁷ T. Hara,³⁰ K. Hayasaka,²⁰ H. Hayashii,²¹ M. Hazumi,⁷ D. Heffernan,³⁰ Y. Hoshi,³⁸ W.-S. Hou,²⁴ Y. B. Hsiung,²⁴ H. J. Hyun,¹⁵ K. Inami,²⁰ A. Ishikawa,³² H. Ishino,⁴¹ R. Itoh,⁷ M. Iwasaki,⁴⁰ D. H. Kah,¹⁵ H. Kaji,²⁰ P. Kapusta,²⁵ N. Katayama,⁷ T. Kawasaki,²⁷ H. Kichimi,⁷ H. J. Kim,¹⁵ Y. J. Kim,⁴ K. Kinoshita,² S. Korpar,^{18,12} Y. Kozakai,²⁰ P. Križan,^{17,12} P. Krokovny,⁷ R. Kumar,³¹ C. C. Kuo,²² Y. Kuroki,³⁰ Y.-J. Kwon,⁴⁵ J. S. Lee,³⁵ M. J. Lee,³⁴ S. E. Lee,³⁴ T. Lesiak,²⁵ J. Li,⁶ C. Liu,³³ D. Liventsev,¹¹ F. Mandl,⁹ A. Matyja,²⁵ S. McOnie,³⁶ T. Medvedeva,¹¹ H. Miyake,³⁰ H. Miyata,²⁷ Y. Miyazaki,²⁰ R. Mizuk,¹¹ M. Nakao,⁷ H. Nakazawa,²² S. Nishida,⁷ O. Nitoh,⁴³ T. Nozaki,⁷ S. Ogawa,³⁷ T. Ohshima,²⁰ S. Okuno,¹³ S. L. Olsen,^{6,8} H. Ozaki,⁷ P. Pakhlov,¹¹ G. Pakhlova,¹¹ H. Palka,²⁵ C. W. Park,³⁵ H. Park,¹⁵ L. S. Peak,³⁶ R. Pestotnik,¹² L. E. Piilonen,⁴⁴ H. Sahoo,⁶ Y. Sakai,⁷ O. Schneider,¹⁶ J. Schümann,⁷ A. J. Schwartz,² K. Senyo,²⁰ M. E. Sevier,¹⁹ M. Shapkin,¹⁰ H. Shibuya,³⁷ J.-G. Shiu,²⁴ J. B. Singh,³¹ A. Sokolov,¹⁰ A. Somov,² S. Stanič,²⁸ M. Starič,¹² T. Sumiyoshi,⁴² F. Takasaki,⁷ N. Tamura,²⁷ M. Tanaka,⁷ G. N. Taylor,¹⁹ Y. Teramoto,²⁹ I. Tikhomirov,¹¹ K. Trabelsi,⁷ S. Uehara,⁷ K. Ueno,²⁴ T. Uglov,¹¹ Y. Unno,⁵ S. Uno,⁷ P. Urquijo,¹⁹ G. Varner,⁶ K. E. Varvell,³⁶ K. Vervink,¹⁶ S. Villa,¹⁶ A. Vinokurova,¹ C. C. Wang,²⁴ C. H. Wang,²³ M.-Z. Wang,²⁴ P. Wang,⁸ Y. Watanabe,¹³ R. Wedd,¹⁹ E. Won,¹⁴ B. D. Yabsley,³⁶ H. Yamamoto,³⁹ Y. Yamashita,²⁶ C. Z. Yuan,⁸ C. C. Zhang,⁸ Z. P. Zhang,³³ and A. Zupanc¹²

(The Belle Collaboration)

¹Budker Institute of Nuclear Physics, Novosibirsk

²University of Cincinnati, Cincinnati, Ohio 45221

³Department of Physics, Fu Jen Catholic University, Taipei

⁴The Graduate University for Advanced Studies, Hayama

⁵Hanyang University, Seoul

⁶University of Hawaii, Honolulu, Hawaii 96822

⁷High Energy Accelerator Research Organization (KEK), Tsukuba

⁸Institute of High Energy Physics, Chinese Academy of Sciences, Beijing

⁹Institute of High Energy Physics, Vienna

¹⁰Institute of High Energy Physics, Protvino

¹¹Institute for Theoretical and Experimental Physics, Moscow

¹²J. Stefan Institute, Ljubljana

¹³Kanagawa University, Yokohama

¹⁴Korea University, Seoul

¹⁵Kyungpook National University, Taegu

¹⁶École Polytechnique Fédérale de Lausanne (EPFL), Lausanne

¹⁷Faculty of Mathematics and Physics, University of Ljubljana, Ljubljana

¹⁸University of Maribor, Maribor

¹⁹University of Melbourne, School of Physics, Victoria 3010

²⁰Nagoya University, Nagoya

²¹Nara Women's University, Nara

²²National Central University, Chung-li

²³National United University, Miao Li

²⁴Department of Physics, National Taiwan University, Taipei

²⁵H. Niewodniczanski Institute of Nuclear Physics, Krakow

²⁶Nippon Dental University, Niigata

²⁷Niigata University, Niigata

²⁸University of Nova Gorica, Nova Gorica

²⁹Osaka City University, Osaka

³⁰Osaka University, Osaka

³¹Panjab University, Chandigarh

³²Saga University, Saga

³³University of Science and Technology of China, Hefei

³⁴Seoul National University, Seoul

³⁵Sungkyunkwan University, Suwon

³⁶University of Sydney, Sydney, New South Wales

³⁷Toho University, Funabashi

³⁸Tohoku Gakuin University, Tagajo

³⁹Tohoku University, Sendai

⁴⁰Department of Physics, University of Tokyo, Tokyo

⁴¹Tokyo Institute of Technology, Tokyo

⁴²Tokyo Metropolitan University, Tokyo

⁴³Tokyo University of Agriculture and Technology, Tokyo

⁴⁴Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061

⁴⁵Yonsei University, Seoul

(Dated: April 22, 2008)

A search for mixing in the neutral D meson system has been performed using semileptonic $D^0 \rightarrow K^{(*)-}e^+\nu$ and $D^0 \rightarrow K^{(*)-}\mu^+\nu$ decays. Neutral D mesons from $D^{*+} \rightarrow D^0\pi_s^+$ decays are used and the flavor at production is tagged by the charge of the slow pion. The measurement is performed using 492 fb^{-1} of data recorded by the Belle detector. From the yield of right-sign and wrong-sign decays arising from non-mixed and mixed events, respectively, we measure the ratio of the time-integrated mixing rate to the unmixed rate to be $R_M = (1.3 \pm 2.2 \pm 2.0) \times 10^{-4}$. This corresponds to an upper limit of $R_M < 6.1 \times 10^{-4}$ at the 90% C.L.

PACS numbers: 14.40.Lb, 13.20.Fc, 12.15.Ff

I. INTRODUCTION

The phenomenon of mixing has been well established in the $K^0\text{--}\bar{K}^0$, $B^0\text{--}\bar{B}^0$ and $B_s^0\text{--}\bar{B}_s^0$ systems. Recently, evidence for mixing in the $D^0\text{--}\bar{D}^0$ system has been obtained with a statistical significance of more than three standard deviations for the first time [1, 2]. In addition, several new measurements help constrain the relevant mixing parameters [3, 4]. The parameters used to characterize $D^0\text{--}\bar{D}^0$ mixing are $x = \Delta m/\bar{\Gamma}$ and $y = \Delta\Gamma/2\bar{\Gamma}$, where Δm and $\Delta\Gamma$ are the differences in mass and decay width between the two neutral charmed meson mass eigenstates, and $\bar{\Gamma}$ is the mean decay width. The mixing rate within the Standard Model is expected to be small [5]: the largest predicted values, which include the impact of long distance dynamics, are of the order $|x|, |y| \lesssim 10^{-2}$.

For $x, y \ll 1$ and negligible CP violation, the time-dependent mixing probability for semileptonic D^0 decays has the following form [6]:

$$\mathcal{P}(D^0 \rightarrow \bar{D}^0 \rightarrow X^+\ell^-\bar{\nu}_\ell) \propto R_M t^2 e^{-\Gamma t}, \quad (1)$$

where R_M is the ratio of the time-integrated mixing probability to the time-integrated non-mixing probability:

$$R_M = \frac{\int_0^\infty dt \mathcal{P}(D^0 \rightarrow \bar{D}^0 \rightarrow X^+\ell^-\bar{\nu}_\ell)}{\int_0^\infty dt \mathcal{P}(D^0 \rightarrow X^-\ell^+\nu_\ell)} \approx \frac{x^2 + y^2}{2}. \quad (2)$$

The mixing rate R_M can be measured directly by using semileptonic decays of D^0 mesons. The most stringent constraint from semileptonic decays, $R_M < 1.0 \times 10^{-3}$ at the 90% confidence level, comes from our previous measurement [7]. Other measurements of R_M using semileptonic decays are less sensitive [8, 9, 10], whereas results from hadronic decays are more precise [11, 12, 13]. In this paper we present an improved search for $D^0\text{--}\bar{D}^0$ mixing using semileptonic decays of charmed mesons, which supersedes our previous measurement [7]. We measure R_M in a 492 fb^{-1} data sample recorded by the Belle detector at the KEKB asymmetric-energy e^+e^- collider [14], at a center-of-mass (cms) energy of 10.58 GeV . The Belle

detector [15] is a large-solid-angle magnetic spectrometer that consists of a silicon vertex detector (SVD), a 50-layer central drift chamber (CDC), an array of aerogel threshold Cherenkov counters (ACC), a barrel-like arrangement of time-of-flight scintillation counters (TOF), and an electromagnetic calorimeter (ECL) comprised of CsI(Tl) crystals located inside a superconducting solenoid coil that provides a 1.5 T magnetic field. An iron flux-return located outside of the coil is instrumented to detect K_L^0 mesons and to identify muons (KLM). Two different inner detector configurations were used. The first 140 fb^{-1} of data were taken using a 2.0 cm radius beam-pipe and a 3-layer silicon vertex detector (SVD-1), and the subsequent 352 fb^{-1} were taken using a 1.5 cm radius beam-pipe, a 4-layer silicon detector (SVD-2) and a small-cell inner drift chamber [16].

To study signal and background distributions we use Monte Carlo (MC) simulated samples [17] in which the number of selected events is about 2.7 times larger than in the data sample.

II. RECONSTRUCTION OF D^0 DECAYS

We select D^0 mesons arising from $D^{*+} \rightarrow D^0\pi_s^+$ decays and reconstruct them as $D^0 \rightarrow K^-\ell^+\nu_\ell$, where ℓ^+ can be either an electron or muon [18]. The notation π_s^+ denotes a *slow* pion, i.e., the pion that originates from the D^{*+} . The average momentum of this pion is only about $0.23\text{ GeV}/c$, whereas the average momentum of the lepton and kaon from the signal decay are $0.96\text{ GeV}/c$ and $1.52\text{ GeV}/c$, respectively. The momenta given in this paper are measured in the laboratory frame, unless otherwise stated; momenta measured in the cms frame are denoted with an asterisk, e.g., p^* . The reconstruction of D^0 mesons in this specific decay chain enables tagging of the D^0/\bar{D}^0 meson flavor at production using the charge of the slow pion π_s^\pm .

There are three detected particles in the final state: π_s^+ , K^- and ℓ^+ , where ℓ^+ can be either a muon or an

electron. The non-mixed decay results in a charge combination $\pi_s^+ K^- \ell^+$, which we refer to as the Right-Sign (RS) charge combination. The mixing process results in $\pi_s^+ K^+ \ell^-$, which we refer to as the Wrong-Sign (WS) charge combination, as summarized in Table I.

TABLE I: The definition of the Right-Sign (RS) and Wrong-Sign (WS) charge combinations.

charge combination	process	name
π_s^+, K^-, ℓ^+	non-mixed	Right-Sign, RS
π_s^+, K^+, ℓ^-	mixed	Wrong-Sign, WS

Because the neutrino is not directly reconstructed, the masses of the D^0 and D^{*+} candidates are smeared. However, by calculating the difference between the two masses, the uncertainty due to the neutrino four momentum cancels to a large extent. Thus

$$\Delta M \equiv M(\pi_s K \ell \nu) - M(K \ell \nu), \quad (3)$$

the reconstructed invariant mass difference between the D^{*+} and the D^0 meson, is the most appropriate observable to extract the number of signal events. For signal events, the distribution of ΔM peaks at $0.145 \text{ GeV}/c^2$, the mass difference between the D^{*+} and D^0 meson (see Fig. 1).

A. Selection criteria

Among all the different processes occurring in e^+e^- collisions, hadronic final states are selected with an efficiency above 99%. The selection is based on the energy of the charged tracks and neutral clusters, total visible energy in the cms system, the z component (opposite to the positron beam direction) of the total cms momentum, and the position of the reconstructed event vertex [19].

Using MC simulation, the criteria to select the signal decays are optimized to give the best significance for the extracted number of mixed (WS) events, $N_{\text{WS}}^{\text{sig}}/\sigma_{N_{\text{WS}}^{\text{bkg}}}$. The uncertainty $\sigma_{N_{\text{WS}}^{\text{bkg}}}$ is due to the fluctuation of the background in the region $\Delta M < 0.16 \text{ GeV}/c^2$; the fluctuation of the signal events at the rate of our previously measured upper limit [7] is negligible. Hence we maximize

$$\frac{N_{\text{WS}}^{\text{sig}}}{\sqrt{N_{\text{WS}}^{\text{bkg}, \Delta M < 0.16}}}. \quad (4)$$

Since the kinematic properties of mixed and non-mixed events are the same, in the optimization the RS signal is used instead of the WS signal. The optimal values for the selection criteria in some observables are correlated and hence the final criteria are obtained by iterative optimization. In the optimization, selection criteria based on

the D^0 proper decay time are also included, as described in Sec. III.

We suppress D^0 mesons arising from $\Upsilon(4S) \rightarrow B\bar{B}$ events in order to avoid the situation in which the selected sample would be composed of two subsamples with different kinematic properties. These $B\bar{B}$ events have different kinematic properties from the decays of D^0 mesons produced in $e^+e^- \rightarrow c\bar{c}$ (continuum events), and a different apparent decay length between the interaction point and the D^0 decay vertex, because of the finite B lifetime. As a result, the D^0 mesons from this source have slightly different resolutions in kinematic variables, and their proper decay time cannot be measured in the same way as for D^0 mesons from the continuum. Since the fraction of D^0 mesons from B decays is smaller than that from the continuum production, and the background contribution from B decays is large, the sensitivity to mixed events is not reduced by rejecting candidates from B decays.

The quantity used to discriminate between $B\bar{B}$ events (spherical) and continuum events (jet-like) is the ratio of the second to zeroth Fox-Wolfram moment, R_2 [20]. To suppress candidates from B decays we demand $R_2 > 0.2$. A further effective rejection of $B\bar{B}$ events is described below.

Tracks with an impact parameter with respect to the interaction point in the radial direction, $dr < 1 \text{ cm}$, and in the beam direction, $|dz| < 2 \text{ cm}$, are considered as π_s^\pm candidates. These criteria remove badly reconstructed tracks and tracks not arising from the interaction point. A slow pion candidate is required to have a momentum smaller than $600 \text{ MeV}/c$. To reduce the background from electrons, we require the electron identification likelihood (based on the information from the CDC, ACC and ECL [21]) of a π_s candidate to be $\mathcal{L}_e < 0.1$, which selects slow pions with an efficiency of 96% and rejects 72% of electrons. The total efficiency of the slow pion selection criteria and tracking is 51%.

Electron candidates are required to have momenta greater than $250 \text{ MeV}/c$ and an electron identification likelihood $\mathcal{L}_e > 0.95$. The efficiency of the identification criterion is 76%; in total around 46% of all generated signal electrons are retained. Muon candidates are required to have momentum greater than $650 \text{ MeV}/c$ and the muon identification likelihood (based on information from the KLM and properties of the track [22]) $\mathcal{L}_\mu > 0.97$; the latter criterion selects muons in the chosen momentum range with an efficiency of 67%. These two requirements are highly correlated since the identification efficiency of muons with momenta lower than $600 \text{ MeV}/c$ is very poor. The efficiency of the above selection criteria and tracking is 30%.

Kaon candidates are required not to satisfy the lepton selection criteria. Kaons from $D^0 \rightarrow K e \nu$ decays should have $p > 850 \text{ MeV}/c$ and kaons from $D^0 \rightarrow K \mu \nu$ decays $p > 600 \text{ MeV}/c$. The difference in this requirement is due to the correlation between lepton momenta and kaon momenta that enters through other kinematic variables,

and due to different background contributions in both decay modes. A combined likelihood for a given track to be a K^\pm , π^\pm or p^\pm is obtained based on the information from the TOF, CDC and ACC [19]. Kaon candidates are selected using $\frac{\mathcal{L}(K)}{\mathcal{L}(K)+\mathcal{L}(\pi)} > 0.51$ (efficiency of 87% for signal kaons in the selected momentum range) and $\frac{\mathcal{L}(K)}{\mathcal{L}(K)+\mathcal{L}(p)} > 0.01$ (efficiency of 99% for signal kaons). Around 42% of all generated kaons in the electron decay mode, and around 48% in the muon decay mode pass the selection criteria.

At this stage, about 17.0% of all generated $D^{*+} \rightarrow \pi_s^+ D^0$, $D^0 \rightarrow K^- e^+ \nu_e$ decays, and about 12.5% of all generated $D^{*+} \rightarrow \pi_s^+ D^0$, $D^0 \rightarrow K^- \mu^+ \nu_\mu$ decays, are reconstructed. Further criteria are applied to improve the sensitivity to mixed events. In the following, these criteria are described and for each of them the signal loss and the background rejection factors are given.

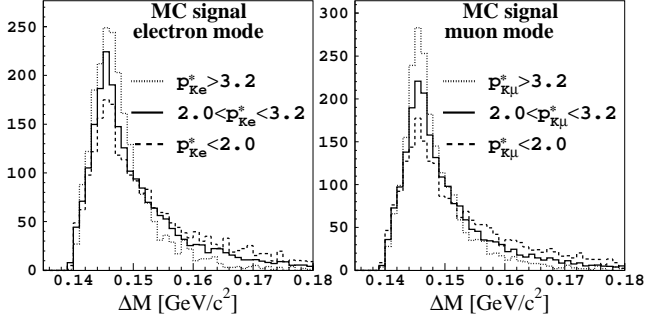


FIG. 1: ΔM distribution for MC-simulated signal events in different $p_{K\ell}^*$ bins: $p_{K\ell}^* < 2.0$ GeV/c (dashed line), 2.0 GeV/c $< p_{K\ell}^* < 3.2$ GeV/c (solid line) and $p_{K\ell}^* > 3.2$ GeV/c (dotted line). The histograms are normalized to the same area. The resolution is improved at higher values of momentum. The left plot is for the electron decay mode and the right one for the muon decay mode.

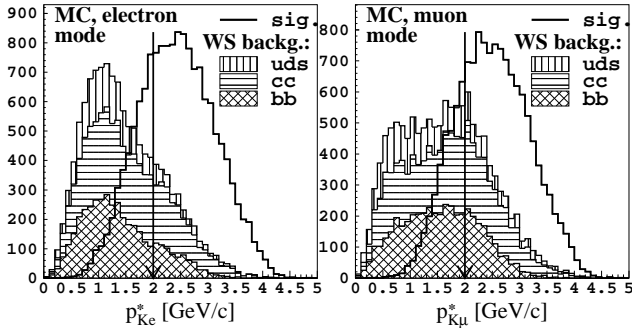


FIG. 2: MC simulated distribution of $p_{K\ell}^*$ for signal (solid line) and background events ($b\bar{b}$, $c\bar{c}$ and uds components of background are shown). The arrow shows the value of the $p_{K\ell}^*$ requirement.

The most effective requirement is the one on the sum of the kaon and lepton momenta, calculated in the cms system, $p_{K\ell}^*$, see Fig. 2. Its optimized value is between 1.7

and 1.9 GeV/c. However, simulated signal events show a clear improvement in the ΔM resolution at higher $p_{K\ell}^*$ values, see Fig. 1. The full width at half maximum (FWHM) of the simulated ΔM distribution reduces from 8.6 MeV/c for $p_{K\ell}^* < 2.0$ GeV/c to 6.9 MeV/c for $p_{K\ell}^* > 3.2$ GeV/c in the electron decay mode, and from 7.7 MeV/c to 5.2 MeV/c for the same $p_{K\ell}^*$ intervals in the muon decay mode. Hence $p_{K\ell}^*$ is required to be at least 2.0 GeV/c, a value that also eliminates a large fraction of D^0 meson decays arising from $\Upsilon(4S) \rightarrow B\bar{B}$ events. This requirement results in a signal loss of 28% in the electron decay mode and 23% in the muon decay mode, while rejecting 76% of the total background in the electron decay mode and 67% of the total background in the muon decay mode.

We apply a selection on the invariant mass of the kaon-lepton system. For the electron decay mode the optimal range is 0.9 GeV/c² $< M(K\ell) < 1.75$ GeV/c² and for the muon decay mode 1.0 GeV/c² $< M(K\mu) < 1.75$ GeV/c², see Fig 3. In the electron decay mode this requirement rejects 25% of the total background at a cost of losing 5.5% of signal events. In the muon decay mode the signal loss is higher, 12%, but so is the background rejection, 44%.

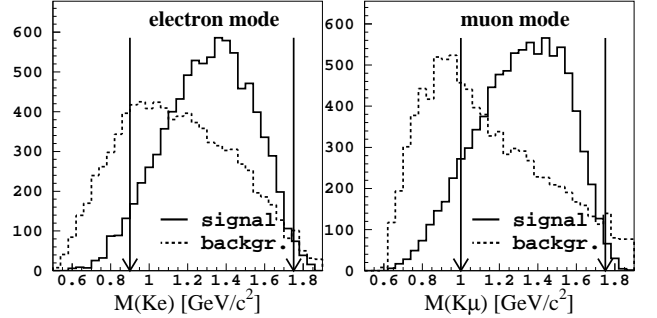


FIG. 3: The distribution of $M(K\ell)$ for signal (solid line) and for background events (dashed), normalized to the same number of entries. The arrows show the values of the selection criteria. The left plot is for the electron decay mode and the right one for the muon decay mode.

D^0 decays to two mesons in the final state are an important source of background. There are four such decays (see Table II). Their branching fractions are at least one order-of-magnitude larger than the effective branching fraction for $D^0 \rightarrow K^+ \ell^- \bar{\nu}_\ell$. Although particle identification reduces their presence in the final sample, it is still important to suppress this background, because its ΔM distribution has a shape similar to that of the signal. Background exhibiting a peak around 0.145 GeV/c² is called *peaking background* and reduces the sensitivity to mixed events much more than the non-peaking background.

In the electron decay mode, the requirement on $M(K\ell)$ rejects 90% of the background arising from the doubly Cabibbo suppressed decay $D^0 \rightarrow K^+ \pi^-$, through misidentification of the pion as an electron; in the muon

TABLE II: Two-body decays, representing a source of a peaking background. The symbol “ \Rightarrow ” represents a misidentification.

decay mode	Br [10^{-3}] [23]	contribution to WS
$D^0 \rightarrow K^- \pi^+$	38.0 ± 0.7	$K^- \Rightarrow \ell^-, \pi^+ \Rightarrow K^+$
$D^0 \rightarrow K^+ \pi^-$	0.143 ± 0.004	$\pi^- \Rightarrow \ell^-$
$D^0 \rightarrow K^- K^+$	3.84 ± 0.10	$K^- \Rightarrow \ell^-$
$D^0 \rightarrow \pi^- \pi^+$	1.36 ± 0.03	$\pi^- \Rightarrow \ell^-, \pi^+ \Rightarrow K^+$

decay mode the suppression rate for this background is 98%. In both decay modes it completely eliminates the background due to misidentification of both pions from $D^0 \rightarrow \pi^- \pi^+$.

The Cabibbo-favoured decays $D^0 \rightarrow K^- \pi^+$ contribute to the WS background if the kaon is misidentified as a lepton and the pion is misidentified as a kaon. To suppress this type of background, the invariant mass of the kaon-lepton system, $M_{\pi K}(K\ell)$, is calculated with the pion mass assigned to the kaon candidate and the kaon mass assigned to the lepton candidate. If $|M_{\pi K}(Ke) - m_{D^0}| < 10 \text{ MeV}/c^2$ in the electron decay mode, and $|M_{\pi K}(K\mu) - m_{D^0}| < 15 \text{ MeV}/c^2$ in the muon decay mode, the $K - \ell$ candidate is rejected. Here m_{D^0} is the mass of the D^0 meson, $1.8645 \text{ GeV}/c^2$ [23]. In the electron decay mode, this requirement rejects $(64 \pm 2)\%$ of the WS background from $D^0 \rightarrow K^- \pi^+$ decays, and $(85 \pm 3)\%$ in the muon decay mode.

To suppress the contribution from $D^0 \rightarrow K^+ K^-$ decays (K^- being misidentified as a lepton), the invariant mass of the kaon-lepton system, $M_{KK}(K\ell)$, is calculated with the kaon mass assigned to both candidates. If $|M_{KK}(Ke) - m_{D^0}| < 10 \text{ MeV}/c^2$ in the electron decay mode, and $|M_{KK}(K\mu) - m_{D^0}| < 15 \text{ MeV}/c^2$ in the muon decay mode, the $K - \ell$ candidate is rejected. In the electron decay mode, this requirement rejects $(70 \pm 5)\%$ of the WS background from $D^0 \rightarrow K^- K^+$ decays, and in the muon decay mode $(89 \pm 1)\%$.

The requirements on $M_{\pi K}(K\ell)$ and $M_{KK}(K\ell)$ result in a signal loss of 3% in the electron decay mode and 2% in the muon decay mode. The rejection of the total background in both decay modes is similar to the signal loss.

B. Rejection of $\gamma \rightarrow e^+ e^-$

An important source of background is due to electrons from photon conversions: either the electron candidate, the slow pion candidate, or both, may be due to $\gamma \rightarrow e^+ e^-$ tracks.

In the electron WS sample, both the slow pion and the signal electron candidates can come from $\gamma \rightarrow e^+ e^-$, and such events tend to have low ΔM values (Fig. 4, left). To suppress this background we calculate $M_{ee}(\pi_s e_{\text{signal}})$, the invariant mass of the $\pi_s - e_{\text{signal}}$ system with the elec-

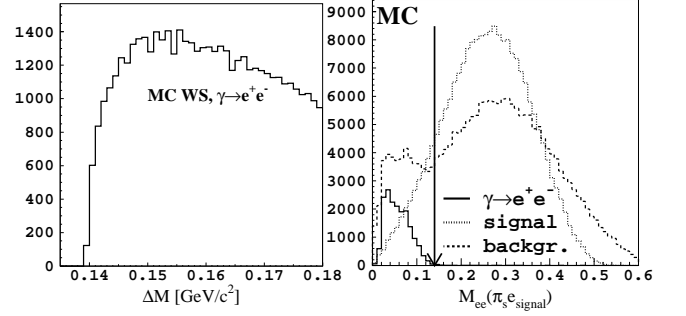


FIG. 4: Left: the ΔM distribution of $\gamma \rightarrow e^+ e^-$ events, where one of the electrons is selected as the slow pion candidate and the other as the electron candidate, before applying the selection on $M_{ee}(\pi_s e_{\text{signal}})$. Right: the $M_{ee}(\pi_s e_{\text{signal}})$ distribution for these events (solid line), for signal events (dotted line) and for the total WS background (dashed line). The arrow shows the value of the selection criterion.

tron mass assigned to both tracks. We reject candidates with $M_{ee}(\pi_s e_{\text{signal}}) < 0.14 \text{ GeV}/c^2$. The requirement rejects more than 99% of this background, see Fig. 4, right. To retain equal reconstruction efficiencies for the mixed and non-mixed events, this requirement is implemented in both RS and WS samples.

Background from $\eta/\pi^0 \rightarrow \gamma\gamma \rightarrow e^+ e^- e^+ e^-$, where one electron from the first photon and one electron from the other photon are taken as the electron and slow pion candidate, exhibits similar behavior to the case where the electrons are both from the same photon. This background is also successfully eliminated by the above selection.

Assuming that the signal electron candidate comes from $\gamma \rightarrow e^+ e^-$, in the electron decay mode we perform a search for the other electron e_2 among all the other tracks in the event with the opposite charge to the signal electron candidate. If $M(e_{\text{signal}} e_2)$, the mass of the $e_{\text{signal}} - e_2$ system, is below $80 \text{ MeV}/c^2$, the electron candidate is rejected.

Assuming that the slow pion candidate is a misidentified electron from $\gamma \rightarrow e^+ e^-$, we perform a search for the other electron (e_2) among all the other tracks in the event. The other electron should have the opposite charge to the slow pion candidate, and an electron likelihood $\mathcal{L}_e > 0.8$ to reduce rejection of true signal slow pions. If $M_{ee}(\pi_s e_2)$, the mass of the $\pi_s - e_2$ system with the electron mass assigned to both tracks, is below $80 \text{ MeV}/c^2$, the slow pion candidate is rejected. This photon conversion rejection is performed for slow pion candidates in both the electron and muon decay modes and results in around 0.4% signal loss and rejects around 2.4% of the total background.

In total, the rejection of photon conversion in the electron decay mode results in a 14% signal loss and 32% rejection of the total WS background.

C. Neutrino reconstruction

Four-momentum conservation in e^+e^- collision implies

$$P_\nu = P_{\text{cms}} - P_{K\ell} - P_{\text{rest}} \quad (5)$$

for the signal decay, where P_{cms} stands for the cms four-momentum of the e^+e^- system and P_{rest} indicates the four-momentum of all detected particles except the charged kaon and the lepton candidates [24]. Eq. (5) is true if all the particles produced in the e^+e^- collision are detected. As the Belle detector covers nearly the entire solid angle around the interaction point, neutrino reconstruction can be successfully performed.

The variable P_{rest} is calculated using all the remaining charged tracks (except the kaon and lepton candidates) with $dr < 2$ cm and $|dz| < 5$ cm, and photons with an energy above 70 MeV. Mass is assigned to a track according to the following criteria:

- A track is assigned the electron mass if its electron likelihood is $\mathcal{L}_e > 0.9$.
- A track is assigned the muon mass if $\mathcal{L}_e < 0.9$ and its muon likelihood is $\mathcal{L}_\mu > 0.9$.
- A track is assigned the kaon mass if $\mathcal{L}_e < 0.9$, $\mathcal{L}_\mu < 0.9$ and $\frac{\mathcal{L}(K)}{\mathcal{L}(K)+\mathcal{L}(\pi)} > 0.5$.
- A track is assigned the proton mass if $\mathcal{L}_e < 0.9$, $\mathcal{L}_\mu < 0.9$, $\frac{\mathcal{L}(K)}{\mathcal{L}(K)+\mathcal{L}(\pi)} < 0.5$ and $\frac{\mathcal{L}(p)}{\mathcal{L}(p)+\mathcal{L}(\pi)} > 0.5$.
- In all other cases the track is assigned the charged pion mass [25].

A first approximation for the neutrino four-momentum P_ν is obtained using Eq. (5) and the resulting ΔM distribution for signal events is shown in Fig. 5 (left) with the dashed line. It peaks at around $0.148 \text{ GeV}/c^2$, a value close to the $D^{*+} - D^0$ mass difference, $0.145 \text{ GeV}/c^2$, and has a FWHM of $58 \text{ MeV}/c^2$.

Two kinematic constraints are used to improve the resolution on the neutrino momentum. To simplify the expressions, we performed the calculation in the cms system, since $\vec{p}_{\text{cms}}^* \equiv 0$. First, the squared invariant mass of the selected particles is calculated using $M^2(K\ell\nu) = (P_\nu^* + P_{K\ell}^*)^2/c^2$. The distribution of $M^2(K\ell\nu)$ is shown in Fig. 6, left. For signal events, the invariant mass should equal m_{D^0} . To reject poorly reconstructed events, exhibiting a large FWHM of the final ΔM distribution, only candidates with $-25 \text{ GeV}^2/c^4 < M^2(K\ell\nu) <$

$64 \text{ GeV}^2/c^4$ are retained. For the selected events, P_{rest}^* is rescaled by a factor ξ requiring

$$M^2(K\ell\nu) = (P_{\text{cms}}^* - \xi P_{\text{rest}}^*)^2/c^2 \equiv m_{D^0}^2. \quad (6)$$

The neutrino four-momentum is then recalculated as $P_\nu^* = P_{\text{cms}}^* - P_{K\ell}^* - \xi P_{\text{rest}}^*$, and a corrected $M(\pi_s K\ell\nu)$ is obtained, where $M(K\ell\nu)$ has been forced to equal m_{D^0} . With this correction, the ΔM distribution has a FWHM of $11 \text{ MeV}/c^2$ in the electron decay mode and $10 \text{ MeV}/c^2$ in the muon decay mode; the improvement is shown in Fig. 5 (left). The distribution of the scale factor ξ for events in the finally selected sample is shown in the left plot of Fig. 7. It peaks at around 1.04 for the electron decay mode and 1.06 for the muon decay mode. The average ξ in both decay modes is around 1.3.

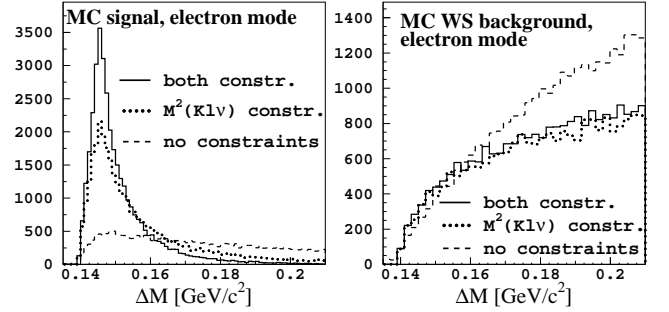


FIG. 5: Distribution of ΔM for signal (left) and background events (right): with the first approximation for the neutrino four-momentum (dashed), after applying the constraint on the D^0 mass (dotted) and with the final neutrino momentum, obtained as described in the text (solid line). Selection criteria on $M^2(K\ell\nu)$ and M_ν^2 have been omitted. The plot is for the electron decay mode; the distributions in the muon decay mode are similar.

As a second kinematic constraint, the square of the missing mass, M_ν^2 , is used. The distribution of M_ν^2 is shown in Fig. 6, right. For events satisfying $-5 \text{ GeV}^2/c^4 < M_\nu^2 < 0.5 \text{ GeV}^2/c^4$, the angle α between the direction of \vec{p}_{rest}^* and the direction of $\vec{p}_{K\ell}^*$ is corrected in order to yield

$$(P_\nu^*)^2 = (P_{\text{cms}}^* - P_{K\ell}^* - \xi P_{\text{rest}}^*)^2 \equiv 0; \quad (7)$$

expressed in terms of energies and magnitudes of three-momenta this yields

$$M_\nu^2 c^4 = (E_{\text{cms}}^* - E_{K\ell}^* - \xi E_{\text{rest}}^*)^2 - p_{K\ell}^{*2} c^2 - \xi^2 p_{\text{rest}}^{*2} c^2 - 2 p_{K\ell}^* \xi p_{\text{rest}}^* c^2 \cos \alpha \equiv 0. \quad (8)$$

The angle α is corrected by rotating \vec{p}_{rest}^* in the plane

determined by the vectors \vec{p}_{rest}^* and $\vec{p}_{K\ell}^*$. The distribu-

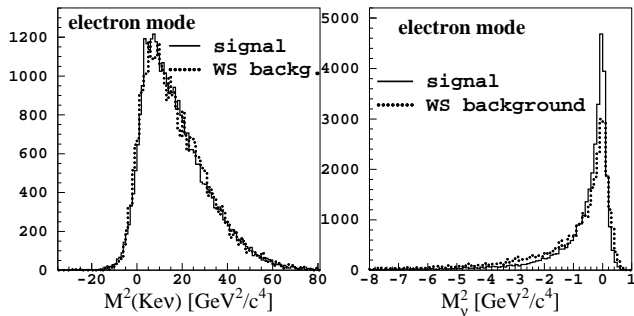


FIG. 6: Left: distribution of $M^2(K\ell\nu)$, calculated using the first approximation neutrino momentum. Right: distribution of M_ν^2 , calculated with the neutrino momentum obtained after applying the constraint on $M^2(K\ell\nu)$. The solid line is for signal and the dotted one for the background. The distributions are shown for the electron decay mode; in the muon decay mode they are similar with a slightly smaller root mean square values.

tion of the correction angle $\alpha_{\text{NEW}} - \alpha_{\text{OLD}}$ for the finally selected signal events is shown in the right plot of Fig. 7. It has a peak at around 1° and an average value of 8° . The final neutrino four-momentum is calculated with the rescaled and rotated P_{rest}^* , using Eq. (5).

The requirements on $M^2(K\ell\nu)$ and M_ν^2 result in a signal loss of 4.5% in the electron decay mode and 4.1% in the muon decay mode while rejecting 9.7% in of background in the electron decay mode and 8.9% of the background in the muon decay mode.

The ΔM distribution obtained using the neutrino four-momentum after the use of kinematic constraints is shown in Fig. 5 (left) as the solid line. The resolution is significantly improved, with the FWHM being about $6.6 \text{ MeV}/c^2$ for the electron decay mode and $6.2 \text{ MeV}/c^2$ for the muon decay mode. Using the MC-simulated background events, it has been verified that such a neutrino reconstruction does not induce any peaking in the background ΔM distribution. From the right plot in Fig. 5 it can be seen that the number of background events in the signal region ($\Delta M < 0.16 \text{ GeV}/c^2$) after applying the constraints only slightly exceeds the number of events without the constraints (by 8.1% in the electron decay mode and 16.6% in the muon decay mode).

For D^{*+} candidates we require $\Delta M < 0.18 \text{ GeV}/c^2$, which retains 97.4% signal events in the electron decay mode and 97.8% in the muon decay mode.

D. D^0 combined with both π_s^+ and π_s^-

After all the above requirements are applied, a small fraction of events contain a D^0 candidate that has been combined with slow pions of opposite charges to form both RS and WS D^* candidates. Events in which such candidates are found are rejected. This veto results in a signal loss of about 2.6% in both electron and muon decay

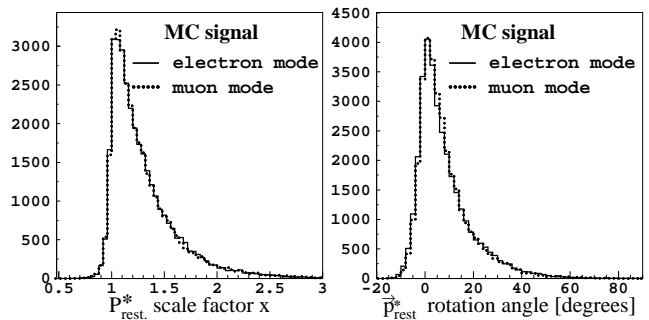


FIG. 7: Left: distribution of the scale factor ξ obtained by the constraint on $M^2(K\ell\nu)$. Right: distribution of the rotation angle for \vec{p}_{rest}^* , obtained by the constraint on M_ν^2 . Both are for signal events; the solid line shows the distributions for electron decays and the dotted one for muon decays.

mode. For the former it rejects 18% of background events and for the latter about 13%.

At this stage the efficiency for reconstructing the signal with $\Delta M < 0.18 \text{ GeV}/c^2$ is found to be $(8.0 \pm 0.3)\%$ for the electron decay mode and $(7.2 \pm 0.3)\%$ for the muon decay mode. The quoted errors include a small variation of efficiency depending on the detector conditions.

A summary of the applied selection criteria is presented in Table III. Since the non-mixed and mixed processes have the same kinematics, the WS and RS efficiencies for all the described selection criteria are the same.

III. PROPER DECAY TIME

As the proper decay time distribution of WS background events tends to have lower values than that of WS signal events, the proper decay time of a D^0 meson can be used to select possible mixed events with a higher purity. Since the information on the proper decay time is used only to increase the sensitivity to WS events, modeling of the proper decay time distribution is not detailed. We do not account for the fact that the associated signal proper decay time resolution function than the signal decays, or for the fact that the resolution depends slightly on the true value of the proper decay time. The effects of these two assumptions were studied carefully; the differences between data and the modeling functions are taken into account in the systematic uncertainties and lead to a negligible change of the final result.

The dimensionless proper decay time (proper decay time in units of $\tau_{D^0} = (410.1 \pm 1.5) \text{ ps}$ [23]) is calculated from the D^0 flight distance l and its momentum \vec{p}_{D^0} :

$$t_{D^0} = \frac{m_{D^0} l}{\tau_{D^0} p_{D^0}}. \quad (9)$$

TABLE III: Summary of all the applied selection criteria.

	electron decay mode	muon decay mode
event	$R_2 > 0.2$	
π_s	impact parameter: $ \delta r < 1 \text{ cm}$, $ \delta z < 2 \text{ cm}$ $p_{\pi_s} < 0.6 \text{ GeV}/c$ electron likelihood < 0.1 $M_{ee}(\pi_s e_2) > 80 \text{ MeV}/c^2$	
ℓ	$p_e > 0.25 \text{ GeV}/c$ electron likelihood > 0.95 $M(e_{\text{signal}} e_2) > 80 \text{ MeV}/c^2$	$p_\mu > 0.65 \text{ GeV}/c$ muon likelihood > 0.97
K	$p_K > 0.85 \text{ GeV}/c$ $\mathcal{L}(K)/[\mathcal{L}(K) + \mathcal{L}(\pi)] > 0.51$ $\mathcal{L}(K)/[\mathcal{L}(K) + \mathcal{L}(p)] > 0.01$	$p_K > 0.6 \text{ GeV}/c$
$K - \ell$	$p_{K\ell}^* > 2.0 \text{ GeV}/c$ $0.9 \text{ GeV}/c^2 < M(Ke) < 1.75 \text{ GeV}/c^2$ $ M_{\pi K}(Ke) - m_{D^0} < 10 \text{ MeV}/c^2$ $ M_{KK}(Ke) - m_{D^0} < 10 \text{ MeV}/c^2$	
$\pi_s - e$	$M_{ee}(\pi_s e_{\text{signal}}) > 0.14 \text{ GeV}/c^2$	/
ν recon.	$-25 \text{ GeV}^2/c^4 < M(K\ell\nu)^2 < 64 \text{ GeV}^2/c^4$ $-5 \text{ GeV}^2/c^4 < P_\nu^2 < 0.5 \text{ GeV}^2/c^4$	
D^{*+}	$\Delta M < 0.18 \text{ GeV}/c^2$ reject D^0 combined with both π_s^+ and π_s^- $1.6 < t_{xy} < 9.0$	

The D^0 momentum is calculated by summing the momenta of the daughter particles. The D^0 flight distance is the distance between the D^0 production vertex, \vec{r}_{prod} , and its decay vertex, \vec{r}_{dec} . The decay vertex is obtained by fitting the kaon and lepton tracks to a common vertex. The production vertex is obtained by extrapolating the D^0 momentum vector to the e^+e^- interaction region. The position and width of this region are determined over a large number of e^+e^- interactions for which the KEKB beam conditions do not change significantly.

According to MC simulation, the resolution on the proper decay time is improved if the D^0 flight distance is calculated as the projection of the $\vec{r}_{\text{dec}} - \vec{r}_{\text{prod}}$ vector on the normalized D^0 momentum vector. Since the interaction region is much narrower in the radial direction, we use only the radial components (x and y) to measure the proper decay time. The radial flight distance l_{xy} is calculated as

$$l_{xy} = \frac{(r_{\text{dec}}^x - r_{\text{prod}}^x, r_{\text{dec}}^y - r_{\text{prod}}^y) \cdot (p_{D^0}^x, p_{D^0}^y)}{\sqrt{(p_{D^0}^x)^2 + (p_{D^0}^y)^2}}. \quad (10)$$

The proper decay time is then evaluated as

$$t_{xy} = \frac{m_{D^0} l_{xy}}{\tau_{D^0} \sqrt{(p_{D^0}^x)^2 + (p_{D^0}^y)^2}}. \quad (11)$$

The observed t_{xy} distribution is smeared due to the experimental resolution. As the data recorded with the SVD-2 configuration has a slightly narrower resolution function than the data taken with the SVD-1 configuration, we perform the measurements of R_M separately for both subsamples. Thus we have four subsamples: the electron subsamples are denoted by e-1 and e-2 for the SVD-1 and SVD-2 configurations, respectively. Similarly the muon subsamples will be denoted as μ -1 and μ -2.

A. Distribution of signal events

We obtain the resolution function for signal events from the data, using the RS decays. To be able to do so, we first determine the shape of the t_{xy} distribution for RS background events, which is shown in Fig. 8, right, as the dashed line. This distribution is also obtained from the data, as described in the following.

The normalized function that describes the t_{xy} distribution of RS background events is

$$\mathcal{F}_{\text{bkg}}^{\text{RS}} = (1 - f_w^b) \cdot \{ [f_e^b E(t; \tau^b) + (1 - f_e^b) \delta(t)] \otimes [f_1^b L(t; b^b) + (1 - f_1^b) L_a(t; b_l^b, b_r^b)] \} + f_w^b \cdot L(t; b_w^b). \quad (12)$$

It is composed as a sum of an exponential function $E(t; \tau^b)$ with the decay time τ^b and a delta function (their fractions are determined by the parameter f_e^b), convolved with a detector resolution. The latter is phenomenologically described by a sum of the Lorentz function $L(t; b^b)$ and an asymmetric Lorentz function $L_a(t; b_l^b, b_r^b)$ (both are explicitly given in the Appendix, Eq. (28) and (29)); b^b , b_l^b and b_r^b are their width parameters and f_1^b determines their fractions in the resolution function. A wide Lorentz function is added to describe the decay times measured from badly reconstructed tracks (outliers); its width parameter is b_w^b and its fraction in the total sample f_w^b . The convolutions are performed numerically by substituting the integral with a sum. It has been verified that the numerical accuracy is satisfactory, *i.e.* not affecting the result.

To determine the eight free parameters of the function from the data, we divide the t_{xy} range into 15 intervals as shown in Fig. 8, left. In each of these 15 intervals we extract the number of RS background events, $N_{\text{RS,bkg}}^i$, by performing fits to the ΔM distribution of RS events. The fit to the ΔM distribution is described in detail in Sec. IV B. The errors include the systematic error due to the finite number of MC simulated events and the uncertainty of the correlated background fraction in the total RS background.

If we divide $N_{\text{RS,bkg}}^i$ in one of the 15 t_{xy} intervals by the total number of background events $N_{\text{RS,bkg}}^{\text{tot}}$ (the sum over the 15 intervals), the obtained fraction is expected to agree with the integral of $\mathcal{F}_{\text{bkg}}^{\text{RS}}$ (Eq. 12) over that t_{xy} interval.

Hence we calculate $N_{\text{RS,bkg}}^i / N_{\text{RS,bkg}}^{\text{tot}}$ in all 15 t_{xy} intervals and determine the eight free parameters of $\mathcal{F}_{\text{bkg}}^{\text{RS}}$ by a χ^2 fit to these fractions. The value of the fitting function in each t_{xy} interval is calculated as the integral of $\mathcal{F}_{\text{bkg}}^{\text{RS}}$ over that t_{xy} interval. The fractions and the result of the fit for the e-2 subsample are shown in Fig. 8, left plot; the obtained background distribution is shown with the dashed line in the right plot. The reduced χ^2 values of the fits are reasonable, ranging from 0.4 to 4 for 7 degrees of freedom. The uncertainties on the fitted parameters of the RS background proper decay time distribution are taken into account when calculating the uncertainty of the result, R_M .

1. RS signal distribution

The proper decay time distribution for RS events is described by

$$\mathcal{F}^{\text{RS}} = \mathcal{N}_{\text{tot}}(f_s \cdot \mathcal{F}_{\text{sig}}^{\text{RS}} + (1 - f_s) \cdot \mathcal{F}_{\text{bkg}}^{\text{RS}}), \quad (13)$$

where \mathcal{N}_{tot} is the total number of RS events and f_s is the signal fraction, obtained from a fit to ΔM in the entire proper decay time region. The f_s values are $(69.9 \pm 0.2)\%$, $(70.8 \pm 0.1)\%$, $(62.7 \pm 0.2)\%$ and $(62.6 \pm 0.1)\%$ for e-1, e-2, μ -1 and μ -2 subsamples, respectively. The function

$\mathcal{F}_{\text{sig}}^{\text{RS}}$ describes the shape of the RS signal events, which is an exponential convolved with the resolution function,

$$\mathcal{F}_{\text{sig}}^{\text{RS}} = E(t; \tau^s) \otimes \mathcal{R}_{\text{sig}}. \quad (14)$$

The resolution function is parameterized as $\mathcal{R}_{\text{sig}} =$

$$f_1^s L_a(t; b_l^s, b_r^s) + f_2^s G(t; \sigma) + (1 - f_1^s - f_2^s) L(t; b_w^s). \quad (15)$$

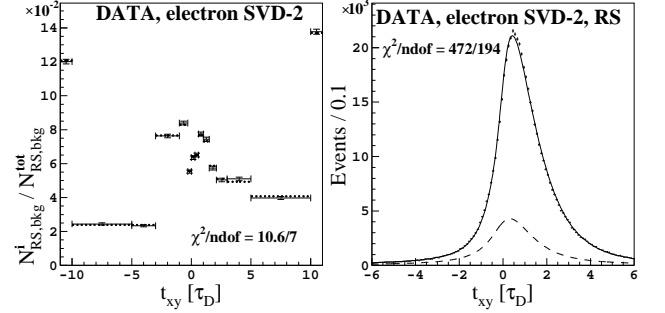


FIG. 8: Left: the fraction of RS background events in 15 proper decay time intervals (error bars) and the result of the fit, described in the text (dotted lines). Note that the t_{xy} intervals are not equidistant, which causes the apparently non-smooth shape of the function. The first and the last interval include events with $t_{xy} < -10$ and $t_{xy} > 10$, respectively. Right: the t_{xy} distribution of RS events (error bars) with the result of the fit (solid line) for the e-2 subsample. The dashed line shows the RS background distribution, $\mathcal{F}_{\text{bkg}}^{\text{RS}}$, obtained from the data as explained in the text.

Here $G(t; \sigma)$ is the Gaussian function. The six free parameters of the resolution function for the RS signal events are f_1^s , f_2^s , b_l^s , b_r^s , σ and b_w^s . They are obtained by a χ^2 fit of \mathcal{F}^{RS} to the proper decay time distribution of RS events. In this fit, the parameters of $\mathcal{F}_{\text{bkg}}^{\text{RS}}$ are fixed to the values previously obtained and τ^s , the dimensionless D^0 decay time, is fixed to 1.0. An example of a fit result is shown in the right plot of Fig. 8. The reduced χ^2 values range between 1.2 and 2.4 for 194 degrees of freedom and exhibit a slight disagreement between the fitting model and the data. The disagreement is accounted for in the systematic error evaluation, resulting in a negligible change of the final result.

2. WS signal distribution

Since decays of the mixed and of the unmixed D^0 mesons have the same kinematic properties, the proper decay time resolution function for both is assumed to be the same. Hence from the RS signal resolution function, \mathcal{R}_{sig} , the proper decay time distribution for WS signal events is calculated:

$$\mathcal{F}_{\text{sig}}^{\text{WS}} = \mathcal{A} t^2 e^{-t/\tau} \otimes \mathcal{R}_{\text{sig}}, \quad (16)$$

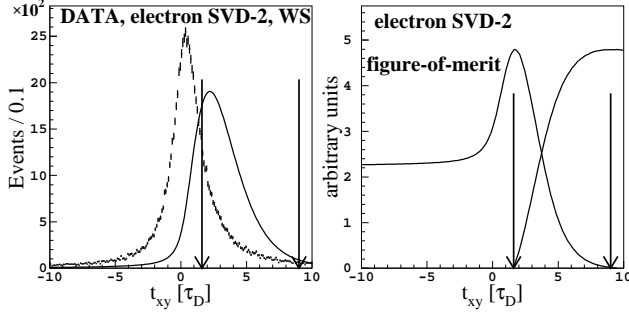


FIG. 9: Left: the proper decay time distribution of the WS data events (error bars) and the distribution for the WS signal events (solid line), obtained from the data according to Eq. (16). The distributions are for the e-2 subsample. Right: the figure-of-merit dependence on the lower and upper limit of the selected t_{xy} interval. When plotting the dependence on the lower limit, the upper limit was set at its optimal value and vice versa. The arrows show the value of the selection criterion, Eq. (17). Distributions for other three subsamples are similar.

where \mathcal{A} is the corresponding normalization constant and τ is fixed to 1.

To select the t_{xy} interval with the highest sensitivity to mixed events, the ratio given in Eq. (4) is maximized. In this optimization we use the calculated distribution for WS signal events, $\mathcal{F}_{\text{sig}}^{\text{WS}}$ (Eq. (16)); for the background we use the distribution of the WS events from the data. Even if the latter contains some mixed events, the effect on the result of the optimization is negligible, since D^0 mixing is small. The distributions of background events, WS signal events and the figure-of-merit can be seen in Fig. (9). The optimal proper decay time intervals for the four subsamples range between 1.6–1.7 and 8.9–9.5. In order to keep the measurement method uniform, we select a common interval for all four subsamples,

$$1.6 < t_{xy} < 9.0. \quad (17)$$

In this interval, about 70% of WS signal events are selected, while rejecting about 80% of background events (the values are similar in each of the four subsamples).

B. Extraction of R_M and further improvement

Because the proper decay time distribution of the RS signal events (e^{-t/τ_D}) is different from that of the WS signal events ($t^2 e^{-t/\tau_D}$), after the application of a selection based on proper decay time, the ratio R_M is obtained as

$$R_M = \frac{N_{\text{WS}}}{N_{\text{RS}}} = \frac{N_{\text{WS}}^i}{N_{\text{RS}}^i} \frac{\epsilon_{\text{RS}}^i}{\epsilon_{\text{WS}}^i}, \quad (18)$$

where $N_{\text{WS,RS}}$ are the numbers of extracted signal events without the t_{xy} selection, and $N_{\text{WS,RS}}^i$ are the numbers of extracted signal events in the selected t_{xy} interval. The

superscript i labels different t_{xy} intervals. The efficiencies ϵ_{RS}^i are obtained by integrating the proper decay time distribution of the RS signal events, $\mathcal{F}_{\text{sig}}^{\text{RS}}$, over the selected t_{xy} interval. Similarly, the efficiencies ϵ_{WS}^i are obtained from the calculated proper decay time distribution of the WS signal events, $\mathcal{F}_{\text{sig}}^{\text{WS}}$. The ratios $\epsilon_{\text{RS}}^i/\epsilon_{\text{WS}}^i$ are listed in Table VII. The errors on the efficiencies quoted there include the uncertainty on the fraction of the RS correlated background (as defined in Sec. IV A), the statistical and systematic uncertainty on the signal fraction in the RS sample, the statistical and systematic uncertainties in the parameters of the RS background and signal t_{xy} distributions and the uncertainties in the world averages of τ_{D^0} and m_{D^0} [23]. The resulting errors on the measured parameter R_M are included in the systematic uncertainty and are negligible. Imperfections in modelling the decay time distributions are included as a separate source of a systematic uncertainty as described below.

To further exploit the proper decay time information, we divide the chosen t_{xy} range (Eq. (17)) into six intervals, with boundaries at 1.6, 2.0, 2.5, 3.1, 4.0, 5.6, and 9.0. The binning is chosen so as to have approximately the same number of events in each interval. The mixing rate is measured in each of the six intervals and the measurements are expected to be consistent. Due to the additional proper decay time information the sensitivity of the final result is expected to be improved in comparison with the sensitivity of a single measurement in the total $1.6 < t_{xy} < 9.0$ range.

IV. SIGNAL YIELD EXTRACTION

According to the MC simulation, the selected sample of RS events includes many candidates from semileptonic decays other than $D^0 \rightarrow K^- \ell^+ \nu$, combined with the correctly reconstructed slow pion. The most important of these decays are:

- $D^0 \rightarrow K^- \pi^0 \ell^+ \nu$,
- $D^0 \rightarrow K^{*-} \ell^+ \nu_\ell$, followed by $K^{*-} \rightarrow K^- \pi^0$,
- $D^0 \rightarrow \pi^- \ell^+ \nu_\ell$,
- $D^0 \rightarrow \rho^- \ell^+ \nu_\ell$, followed by $\rho^- \rightarrow \pi^- \pi^0$,
- $D^0 \rightarrow K^{*-} \ell^+ \nu_\ell$, followed by $K^{*-} \rightarrow \bar{K}^0 \pi^-$.

The final state lepton in these decays is of the same charge as in the decay $D^0 \rightarrow K^- \ell^+ \nu$, so its charge can be used to tag mixing in the same way. In the last three decays with a π^- instead of the K^- in the final state, the pion is misidentified as a kaon. The selected sample also includes candidates where the D^0 semileptonic decay is correctly reconstructed, but the slow pion decays in flight to a muon, $\pi_s^+ \rightarrow \mu^+ \nu_\mu$, and then the muon is misidentified as the slow pion. The muon has the same charge as the slow pion and can be used to tag the D^0 flavor at

production in the same way as the slow pion. Hence all these processes are treated as part of the signal and will be referred to as *associated signal* decays.

According to MC simulation, the associated signal decays have a similar ΔM distribution to the $D^0 \rightarrow K^- \ell^+ \nu$ decays. Due to unreconstructed or misidentified particles in these final states, the FWHM of the distribution is larger, $12.3 \text{ MeV}/c^2$ in the electron decay mode and $14.3 \text{ MeV}/c^2$ in the muon decay mode (see Fig. 10). For the same reasons, the proper decay time distribution is slightly different from that for the signal. The fraction of the associated signal decays in the sample of all reconstructed signal decays can be found in Table IV. There is also a small fraction of signal events (around 1%) from $B\bar{B}$ events. Due to the lower average momentum of D mesons from B decays, the ΔM distribution for signal events from B decays is slightly wider than that for signal events from $c\bar{c}$ events; it is similar to the ΔM distribution of the associated signal decays.

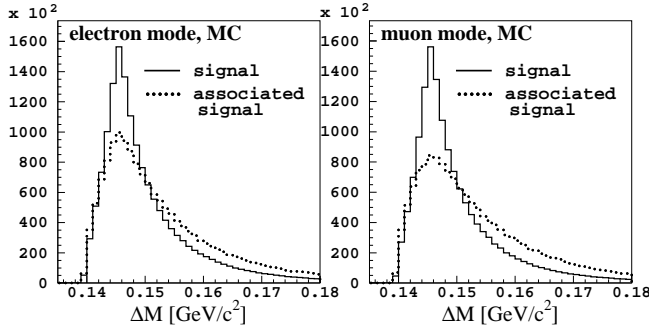


FIG. 10: The simulated ΔM distribution for the signal decays (solid line) and for the associated signal decays (dotted line), for both the (left) electron and (right) muon decay modes.

TABLE IV: The associated signal fraction and the fraction of signal from B decays in the total signal, in [%], as obtained from MC simulation. The quoted uncertainties are statistical only. The fractions are shown for the entire proper decay time interval and for $1.6 < t_{xy} < 9.0$.

	assoc. sig. [%]		sig. from B decays [%]	
	all t_{xy}	$1.6 - 9.0$	all t_{xy}	$1.6 - 9.0$
e	16.58 ± 0.05	17.7 ± 0.1	1.20 ± 0.01	1.18 ± 0.03
μ	11.54 ± 0.04	12.3 ± 0.1	1.07 ± 0.01	1.04 ± 0.03

A. Background

The background is divided into two categories: the correlated background and the uncorrelated background. The ΔM shapes of both background components can be seen in Fig. 11.

Correlated background is the background where the lepton candidate or the kaon candidate, or both, originate from the same decay chain as the slow pion candidate. The angular correlation between the slow pion and D^0 candidates leads to a concentration of events at low values of ΔM .

The remaining, uncorrelated background has a ΔM distribution that rises steadily from threshold, as the available phase space increases. This component is dominant, especially in the WS sample (see Table V). The fraction in the muon decay mode is larger than in the electron decay mode, because the probability for a kaon or a pion to be misidentified as a muon is larger than the probability to be misidentified as an electron. The fraction is larger in the RS sample than in the WS sample due to the larger branching fractions of the Cabibbo favored decays. Selecting the proper decay time interval $1.6 < t_{xy} < 9.0$, decreases the fraction of the correlated background in the total background. According to MC simulation, the correlated background has three components: background from $D^{*+} \rightarrow \pi_s^+ D^0$ decays, which has the largest fraction and is the most strongly peaked of the three, background from $D^{*0} \rightarrow \gamma D^0$, $\gamma \rightarrow e^+ e^-$ decays in which one of the two electrons from γ conversion is taken as a slow pion candidate, and background from $K_S^0 \rightarrow \pi^+ \pi^-$ decays, where one of the pions is taken as the slow pion candidate and the other is assigned as a kaon or lepton candidate.

TABLE V: The fraction of the correlated background in the total background in [%], as obtained from MC simulation, for the electron and muon decay mode, in both RS and WS samples. Fractions for the entire proper decay time interval and for the selected interval are shown. The quoted uncertainties are statistical only.

	Right Sign sample		Wrong Sign sample	
	all t_{xy}	$1.6 - 9.0$	all t_{xy}	$1.6 - 9.0$
e	34.49 ± 0.07	25.33 ± 0.15	7.22 ± 0.05	5.64 ± 0.10
μ	40.07 ± 0.06	39.62 ± 0.14	14.87 ± 0.06	14.76 ± 0.14

The ΔM distribution for the total RS background can be seen in Fig. 12 as the dashed line. To fit the ΔM distribution of the data as explained in Sec. IV B, the ΔM distribution of the RS background events is obtained from MC simulation.

The ΔM distribution for the total WS background is shown in Fig. 11 as the solid histogram. One can see that the difference in shape between the uncorrelated background and the total background is smaller in the electron decay mode than in the muon decay mode. This is both due to the larger correlated background component in the muon decay mode (see Table V) and due to its shape, shown in the same figure.

To check the MC simulation and to avoid systematic errors arising from any discrepancy with the data, the uncorrelated background in the WS sample was described

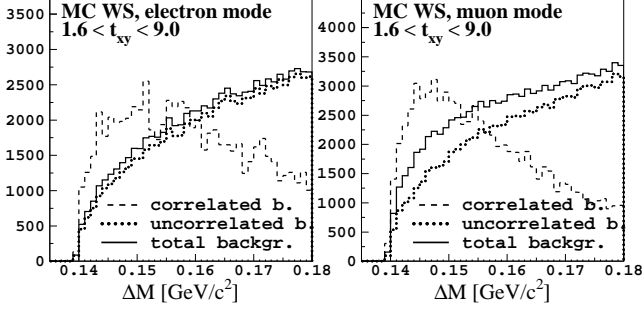


FIG. 11: The simulated ΔM distribution for WS background events. The solid line shows the total background, the dotted line shows the uncorrelated background and the dashed one shows the correlated background, multiplied by a factor of 15 and 5 in the electron and muon samples, respectively. The plots are for the selected proper decay time interval. The left plot is for the electron decay mode and the right one for the muon decay mode.

using the data. This background is modeled by combining slow pion candidates and D^0 candidates from different events. Technically this is done by *embedding* slow pion candidates into other events according to the following procedure:

- All slow pion candidates from an event in which a D^0 candidate was found, are taken to be embedded into other events.
- We denote by \mathcal{N} the number of slow pion candidates which form a WS combination in their original event, and their charge by \mathcal{Q} . Slow pions are embedded only into events with the same value of \mathcal{N} and only the embedded slow pions of charge \mathcal{Q} are used to form the WS combinations.
- For each value of \mathcal{N} , slow pion candidates from \mathcal{N}_A different events are stored to be embedded into other events.
- Each D^0 candidate is combined with slow pions from several other events; slow pions from a maximum of \mathcal{N}_A events are used. Once a combination satisfying all the D^{*+} requirements is obtained (including $\Delta M < 0.18 \text{ GeV}/c^2$), further combinations are not formed.
- Once an embedded slow pion is used to form a combination satisfying all the D^{*+} requirements, none of the remaining slow pions from the same event is embedded into other events.

With these requirements, the ΔM distribution of the sample of embedded slow pions slightly depends on \mathcal{N}_A , the maximum number of events from which the embedded slow pions are taken and tested with a single D^0 candidate. The dependence is due to the ΔM dependence

on the slow pion momentum. Slow pions with higher momenta tend to form D^{*+} candidates with slightly higher ΔM values, hence their probability to form a D^{*+} candidate with $\Delta M < 0.18 \text{ GeV}/c^2$ is smaller. Increasing \mathcal{N}_A enables these slow pions to be tested with a larger number of D^0 candidates and enhances the probability to form a combination with $\Delta M < 0.18 \text{ GeV}/c^2$. Thus increasing \mathcal{N}_A slightly enhances the contribution at higher ΔM values.

The most appropriate \mathcal{N}_A value is determined from the data by observing samples of $t_{xy} < 0.0$. In this proper decay time region, the expected fraction of the mixed signal events in the WS background is much smaller (around 18 times in the electron and around 27 times in the muon decay mode) than in the region $1.6 < t_{xy} < 9.0$. Hence it is safe to assume that the WS data sample of $t_{xy} < 0.0$ contains no mixed signal events. For $t_{xy} < 0.0$ we combine the SVD-1 and SVD-2 subsamples and compare the ΔM distribution of the WS data with the ΔM distribution of the subsample used to describe the WS background, *i.e.* embedded slow pions with the addition of the MC correlated background events (as explained at the end of this section). The ΔM distributions are compared by observing the value of

$$r_{155} = N_{\Delta M < 0.155} / N_{\Delta M < 0.18}, \quad (19)$$

the ratio of the number of events with $\Delta M < 0.155 \text{ GeV}/c^2$ and the number of events with $\Delta M < 0.18 \text{ GeV}/c^2$. This is a representative observable that is used to characterize the ΔM distribution by a single number. The values for the data and for the background with three different values of \mathcal{N}_A are shown in Table VI.

TABLE VI: Comparison of the r_{155} values for the WS data and the modeled WS background with $t_{xy} < 0.0$, using different values of \mathcal{N}_A .

$t_{xy} < 0.0$	electron mode		muon mode	
	\mathcal{N}_A	$r_{155}[\%]$	\mathcal{N}_A	$r_{155}[\%]$
data		25.66 ± 0.18		29.00 ± 0.16
background	40	25.80 ± 0.05	20	29.11 ± 0.05
	45	25.70 ± 0.05	25	29.00 ± 0.05
	50	25.61 ± 0.05	30	28.90 ± 0.05

From Table VI one can see that for the electron decay mode the best agreement between the data and the background ($t_{xy} < 0.0$) is for $\mathcal{N}_A = 45$ and in the muon decay mode for $\mathcal{N}_A = 25$. The dependence of the final result on the \mathcal{N}_A value is taken into account when evaluating the systematic uncertainty.

As a cross-check, we combine SVD-1 and SVD-2 subsamples and compare the ΔM distribution of the MC uncorrelated WS background with that of the embedded slow pions. Their agreement is good.

To obtain the final ΔM distribution of WS background events, the ΔM shape of the WS-correlated background is taken from MC simulation and added to the sample of

the embedded slow pions in the same fraction as found by MC simulation. The uncertainty on this fraction is taken into account when evaluating the systematic uncertainty.

B. Fit to ΔM distribution

To extract the signal yield, we perform a binned maximum likelihood fit to the ΔM distribution, assuming a Poisson distribution of events in ΔM bins and thus maximizing

$$\mathcal{L} = \prod_{j=1}^{N_{\text{bin}}} \frac{e^{-\mu(\Delta M_j)} \cdot (\mu(\Delta M_j))^{N_j}}{N_j!}. \quad (20)$$

Here N_j is the number of entries in the j -th bin and $\mu(\Delta M_j)$ is the expected number of events in this bin, given by

$$\mu(\Delta M_j) = \mathcal{N}_R [f_s P_s(\Delta M_j) + (1 - f_s) P_b(\Delta M_j)]. \quad (21)$$

P_s is the signal ΔM distribution obtained from MC simulation. P_b is the background ΔM distribution composed as described above. The signal fraction f_s is the only free parameter in the fit. \mathcal{N}_R is the number of entries in the fitted histogram. $N_{\text{bin}} = 45$ is the number of intervals in the ΔM distribution. The quoted χ^2 values are obtained using $\chi^2 = \sum_{j=1}^{N_{\text{bin}}} \frac{(N_j - \mu(\Delta M_j))^2}{\sigma_j^2}$, where σ_j includes the statistical uncertainties of the fitting histograms, $\sigma_j^2 = N_j + \sigma_{P_s,j}^2 + \sigma_{P_b,j}^2$.

C. The RS signal yield

The fit to the ΔM distribution in the RS sample is performed as described above; examples of the fit result are shown in Fig. 12. In the total t_{xy} range, the signal fraction f_s is about 70% in the electron decay mode and about 63% in the muon decay mode. The fraction is largest for $1.6 < t_{xy} < 2.0$ (82% in the electron decay mode and 74% in the muon decay mode) and decreases at larger t_{xy} values: for $5.6 < t_{xy} < 9.0$ it is 62% in the electron decay mode and 54% in the muon decay mode. The χ^2 values of the fits in the individual t_{xy} intervals are in good agreement with the expectation for 40 degrees of freedom. In the total t_{xy} range, the reduced χ^2 values are larger than expected (values of 1.5–2.6 for 40 degrees of freedom). This is explained by a difference in the amount of associated signal between the data and the MC simulation. Repeating the fits with a fraction of the associated signal as the second free parameter yields reduced χ^2 values around 1.0 also for the total t_{xy} region. This effect is considered in the estimate of the systematic uncertainty. The numbers of RS signal events are given in Table VII.

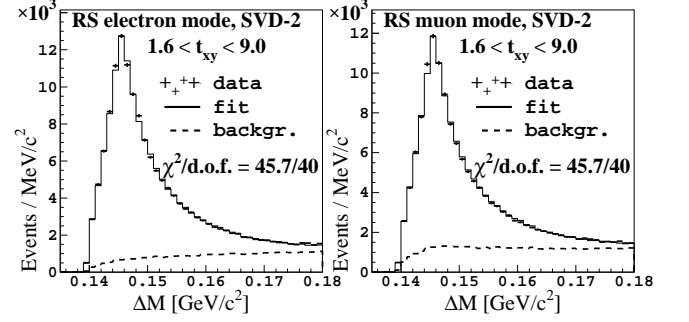


FIG. 12: The ΔM distribution of the RS events for $1.6 < t_{xy} < 9.0$, SVD-2. The dashed line represents the background, the solid line is the result of the fit, described in the text, and the points with error bars are the data. The left plot is for the e-2 subsample, the right one for the μ -2 subsample.

V. RESULT

As the kinematic properties of the RS and WS decays are the same, we use the ΔM shape of the MC simulated RS signal events also for the WS signal decays. The ΔM distribution of the WS background events is obtained as described in Sec. IV A. By fitting the ΔM distribution as described in Sec. IV B, we extract the number of the mixed signal events in the four WS subsamples (e-1, e-2, μ -1, μ -2). The ΔM distributions and the χ^2 values of the fits for all the subsamples and different proper decay time intervals are shown in Figs. 13, 14. The extracted WS signal yields are given in Table VII.

For each of the four subsamples we determine the mixing ratio R_M by three different methods which are discussed below.

- 1) The fit to ΔM in the RS and WS samples is performed without any selection based on the proper decay time measurement. The ratio R_M is calculated as the ratio of the obtained number of WS and RS signal events, $N_{\text{WS}}/N_{\text{RS}}$. The results can be found in Table VII in rows labeled “all t_{xy} ”.
- 2) The fit to ΔM distributions for the RS and WS sample is performed for events with $1.6 < t_{xy} < 9.0$. The ratio R_M is calculated as $N_{\text{WS}}^i/N_{\text{RS}}^i \times \epsilon_{\text{RS}}^i/\epsilon_{\text{WS}}^i$. The results are given in Table VII in rows labeled “1.6–9.0”. The resulting statistical uncertainty of the result is around 34% smaller than the one obtained by method 1).
- 3) The third result, given in rows labeled “combined” of Table VII, is a χ^2 fit of a constant to the six R_M^i values measured in the six proper decay time bins. The six R_M^i values and the result of the fit for each of the four subsamples are shown in Fig. 15. The statistical uncertainty of this result is 2–3% smaller than in method 2), because additional information

on proper decay time is included through the six $\epsilon_{\text{RS}}^i/\epsilon_{\text{WS}}^i$ ratios.

Using the MC simulation, we verified that method 3) has the best sensitivity; we therefore quote as our final result the value obtained by method 3). To illustrate the effect of including the proper decay time information, we show also the results of methods 1) and 2).

From Table VII one can see that the central values obtained by using the three methods, are slightly different. In evaluating the significance of the difference between the result of method 1) and the result of method 2), we have accounted for the ratio of the proper decay time efficiencies and for the statistical correlation between the samples. The differences are within the expected statistical fluctuations: they range between -0.6 and $+1.4$ standard deviations. By using toy MC simulation, it has also been verified that the differences in the central values between methods 2) and 3) are within the range of expected statistical fluctuations. For the default method 3), results for R_M in all four subsamples are consistent with the null value. The χ^2 values, shown in Fig. 15, are in good agreement with the expected χ^2 distribution for 5 d.o.f., which has a maximum at the value of 3.0.

The combined result for the electron decay mode is obtained by a χ^2 fit to the values for the e-1 and e-2 subsamples, obtained by method 3). The fit yields $R_M^e = (-0.6 \pm 2.6) \times 10^{-4}$ with a χ^2 value of 0.1 per 1 degree of freedom. The combined result for the muon decay mode is obtained in the same way; the χ^2 fit yields $R_M^\mu = (5.9 \pm 3.7) \times 10^{-4}$ with a χ^2 value of 0.4.

The combined result, taking into account the statistical uncertainty only, is obtained by a χ^2 fit to the four values (electron and muon decay mode, SVD-1 and SVD-2); it yields a value of

$$R_M^{\text{stat.}} = (1.6 \pm 2.2) \times 10^{-4}, \quad (22)$$

where the quoted uncertainty is statistical only. The χ^2 value is 2.5 for three degrees of freedom. The R_M values for the four subsamples and the result of the fit are shown in the left plot of Fig. 16. To obtain the final result, the partially correlated systematic uncertainties have to be studied and taken into account.

A. Systematic uncertainties

In the following subsection, different sources of the systematic uncertainties are discussed and the systematic uncertainties are given; they are summarized in Table VIII.

1. Finite statistics of the fitting distributions

One of the main sources of systematic uncertainty is the limited statistics of the samples used to obtain the signal and background ΔM distributions used in the

TABLE VII: The number of fitted signal events in the RS and WS samples, the ratio of RS and WS t_{xy} efficiencies, and the resulting R_M^i value for each proper decay time interval for the four subsamples. The results of the fit to the six individual R_M^i values are denoted as “combined”.

t_{xy}	N_{RS}^i	N_{WS}^i	$\epsilon_{\text{RS}}^i/\epsilon_{\text{WS}}^i$	$R_M^i [10^{-4}]$
e-1 subsample:				
1.6–2.0	12578±94	4.8±27.2	0.915±0.007	3.5±19.8
2.0–2.5	11273±89	10.9±26.3	0.634±0.004	6.1±14.8
2.5–3.1	8975±84	14.7±25.6	0.443±0.002	7.2±12.6
3.1–4.0	7937±83	−28.0±25.9	0.310±0.003	−10.9±10.1
4.0–5.6	6394±85	−21.2±28.6	0.223±0.003	−7.4±10.0
5.6–9.0	4196±89	15.9±29.8	0.223±0.003	8.4±15.8
combined				−1.7±5.2
1.6–9.0	51325±213	−11.5±65.4	0.413±0.001	−0.9±5.3
all t_{xy}	183496±443	70.1±141	1	3.8±7.7
e-2 subsample:				
1.6–2.0	32616±150	−19.1±44.0	0.881±0.003	−5.2±11.9
2.0–2.5	28711±146	−11.4±41.7	0.603±0.002	−2.4±8.8
2.5–3.1	22513±131	52.5±41.9	0.415±0.002	9.7±7.7
3.1–4.0	18941±132	−22.6±41.1	0.285±0.002	−3.4±6.2
4.0–5.6	14796±129	−18.6±42.3	0.198±0.002	−2.5±5.7
5.6–9.0	9072±128	25.2±46.5	0.186±0.002	5.2±9.5
combined				−0.1±3.1
1.6–9.0	126539±332	−10.7±102	0.389±0.001	−0.3±3.1
all t_{xy}	469947±701	−369±222	1	−7.8±4.7
μ -1 subsample:				
1.6–2.0	11314±111	14.2±34.7	0.921±0.005	11.6±28.2
2.0–2.5	10185±109	−1.8±33.5	0.637±0.004	−1.1±21.0
2.5–3.1	7893±98	3.5±30.7	0.440±0.003	1.9±17.1
3.1–4.0	6804±96	−5.5±31.8	0.303±0.002	−2.5±14.2
4.0–5.6	5350±97	23.7±33.0	0.214±0.002	9.5±13.2
5.6–9.0	3670±90	−12.8±35.4	0.217±0.003	−7.6±20.9
combined				2.2±7.1
1.6–9.0	45181±245	−11.2±79.9	0.410±0.001	−1.0±7.2
all t_{xy}	163215±485	−204±180	1	−12.5±11.0
μ -2 subsample:				
1.6–2.0	27612±180	71.4±54.8	0.876±0.015	22.7±17.4
2.0–2.5	23695±170	9.3±52.3	0.595±0.010	2.3±13.1
2.5–3.1	18905±154	82.3±49.8	0.405±0.006	17.6±10.7
3.1–4.0	15488±150	51.1±50.1	0.273±0.004	9.0±8.8
4.0–5.6	11989±144	20.4±51.1	0.186±0.007	3.2±7.9
5.6–9.0	7146±138	−20.3±56.5	0.171±0.016	−4.9±13.6
combined				7.4±4.4
1.6–9.0	104556±381	192±125	0.380±0.002	7.0±4.5
all t_{xy}	396151±761	170±284	1	4.3±7.2

ΔM fit to data. To estimate this, we vary the contents of all bins of the RS and WS, signal and background ΔM distributions independently in accordance with each bin’s statistical uncertainty. We repeat the fit to the RS and WS data, calculate the corresponding R_M^i in each proper decay time interval, and obtain a new R_M value. Repeating the procedure 1000 times, the obtained distribution of R_M values has a Gaussian shape. The sigma of

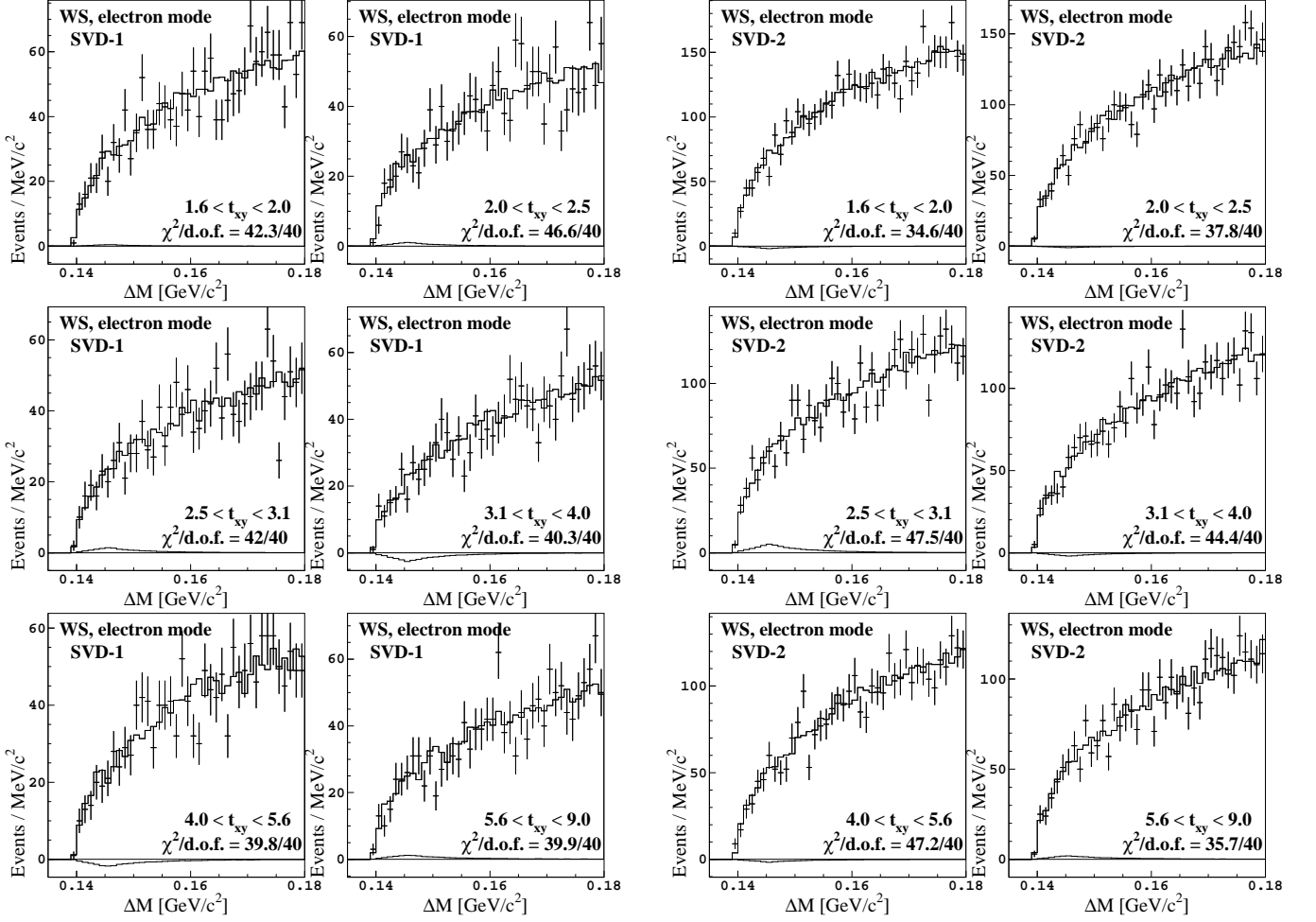


FIG. 13: The ΔM distribution of WS events in the six proper decay time intervals for the e-1 (left) and e-2 (right) subsamples. The points with error bars are the data, the histogram represents the result of the fit, described in the text, and the small contribution on the horizontal axis shows the fitted signal yield.

the Gaussian, fitted to the distribution, is taken as the systematic uncertainty due to the limited statistics of the fitting distributions. The uncertainties are listed in Table VIII, line 1. Larger uncertainties in the muon decay mode reflect the fact that, compared to the electron decay mode, the muon background is larger especially in the signal region, and secondly, the embedded slow pion sample is smaller due to the smaller \mathcal{N}_A value used. Since this uncertainty is statistical in nature, it is considered to be completely uncorrelated between the four subsamples (e-1, e-2, μ -1, μ -2).

2. The amount of WS correlated background

The normalization of the WS-correlated background is determined by MC simulation, taking into account the central values of branching fractions [23], of decay modes that contribute to this background.

From MC simulation studies we find that in the electron decay mode the largest contributions to the WS

correlated background come from the following decays: $D^0 \rightarrow K^- e^+ \nu_e$ (D^0 mesons mainly from $D^{*0} \rightarrow D^0 \gamma$ decays, 34% of the correlated background), $D^0 \rightarrow K^- \pi^+ \pi^0$ (12%), $D^0 \rightarrow K^- e^+ \nu_e \pi^0$ (7%), $D^0 \rightarrow K^- K^+$ (7%). In total 60% of the correlated background comes from these decays.

In the muon decay mode, the largest contributions to the WS correlated background come from $D^0 \rightarrow K^- \pi^+ \pi^0$ (19%), $D^0 \rightarrow K^- \mu^+ \nu_\mu$ (12%), $D^0 \rightarrow K^- \pi^+ \pi^0 \pi^0$ (7%), $D^0 \rightarrow K^- \pi^+ \pi^- \pi^+$ (7%), $D^0 \rightarrow K^+ \pi^- \bar{K}^0$ (4%), $D^0 \rightarrow K^+ K^- \bar{K}^0$ (4%). In total 53% of the correlated background comes from these decays. We calculate the weighted average of the relative uncertainties of the branching fractions [23] for the stated decay modes. For the electron decay mode, the averaged relative uncertainty is $\pm 3.6\%$ and for the muon decay mode $\pm 5.9\%$. To take into account the uncertainties of the branching fractions used in the MC simulation, we repeat the WS fits, changing the amount of the total WS correlated background by the average uncertainties on the branching fractions. The differences between the re-

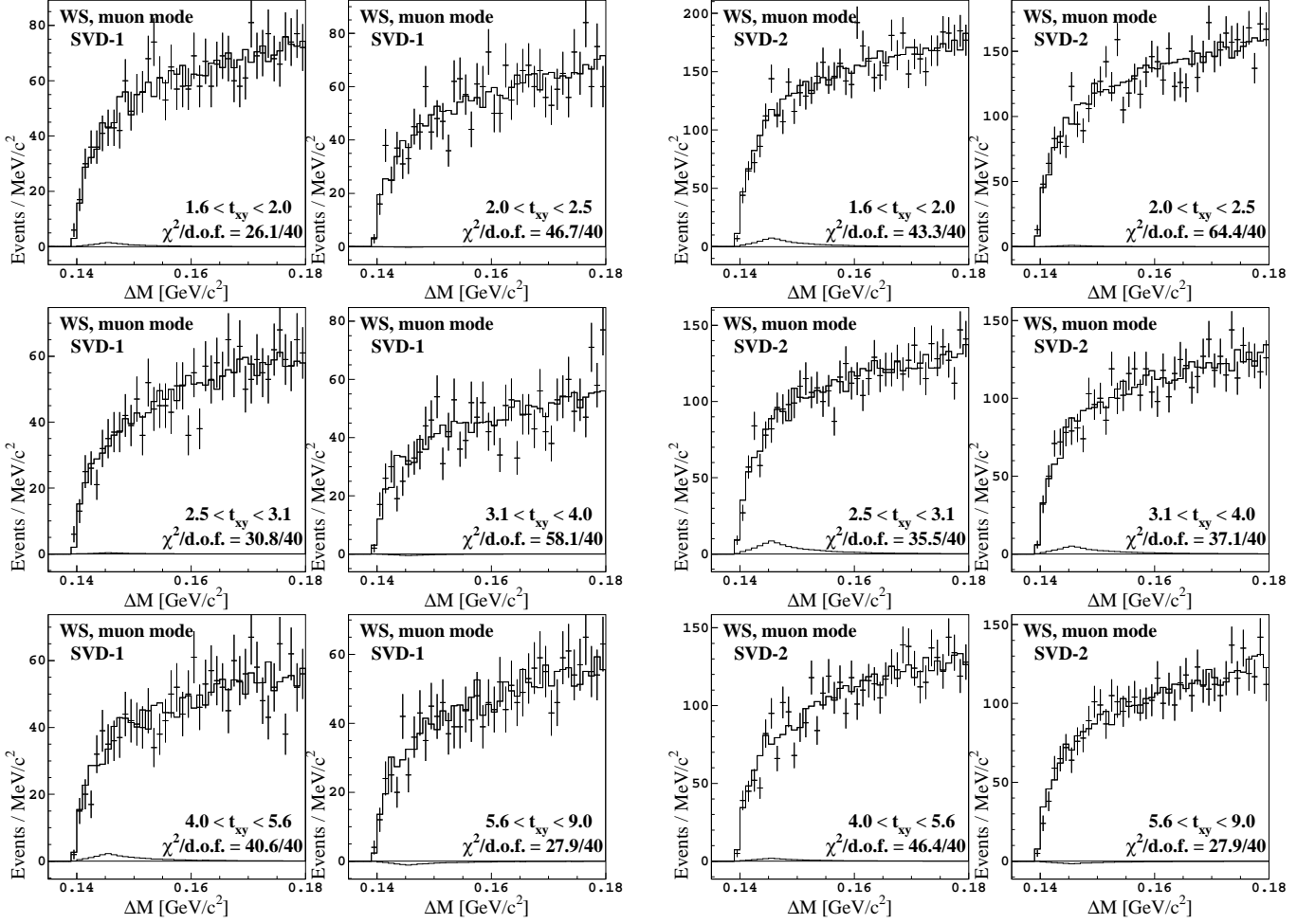


FIG. 14: Same as Fig. 13 for the μ -1 and μ -2 subsamples.

sulting R_M values and the default values are taken as the systematic uncertainty from this source; they are listed in Table VIII, line 2.

This procedure is conservative for two reasons. First, by varying the total correlated background instead of varying its individual components, the uncertainties on the branching fractions are implicitly considered to be 100% correlated, resulting in the maximum possible systematic uncertainty. Second, the modes comprising the correlated background contribute significantly also to the uncorrelated background. Taking this into account would lead to a smaller change in R_M .

This uncertainty is larger in the muon decay mode, because the probabilities to misidentify a pion or kaon as a muon are much larger than the corresponding probabilities for misidentification as an electron. Consequently, in the muon decay mode the fraction of the correlated background is significantly larger (Table V), its ΔM shape tends to lower values (Fig. 11) and its averaged uncertainty of the branching fractions is larger.

The systematic uncertainty from this source is the

same for the SVD-1 and SVD-2 subsample. Since a significant part of the correlated background is due to decays common to the electron and the muon decay modes, the systematic uncertainties for both decay modes are highly correlated. Hence the systematic uncertainty from this source will be treated as 100% correlated for all four subsamples (e-1, e-2, μ -1, μ -2).

3. The ΔM shape of the WS uncorrelated background

We also conservatively account for the uncertainty of the ΔM shape of the WS uncorrelated background. We vary \mathcal{N}_A within the limits given by the statistical uncertainties of the r_{155} values in Table VI. The r_{155} statistical uncertainty for the data is ± 0.18 in the electron decay mode and ± 0.16 in the muon decay mode. For the embedded slow pion sample, the r_{155} value changes by 0.1 for $\Delta \mathcal{N}_A = 5$. Hence \mathcal{N}_A is varied by ± 9 in the electron decay mode, and by ± 8 in the muon decay mode. With the new ΔM distributions we repeat the fit to WS data,

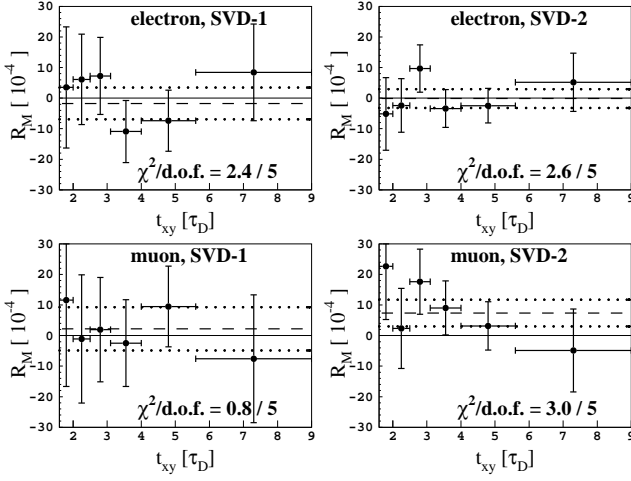


FIG. 15: The resulting R_M^i values for the four subsamples and their average value (dashed line). The dotted lines represent the $\pm 1\sigma$ interval. The solid line corresponds to no mixing.

recalculate the R_M values and quote the differences from the default values as the systematic uncertainties from this source. These uncertainties are listed on line 3 of Table VIII.

Since \mathcal{N}_A is determined for the electron and muon decay mode separately and the uncertainty on \mathcal{N}_A is statistical in nature, this systematic uncertainty is considered to be completely uncorrelated between both decay modes. On the other hand, \mathcal{N}_A is determined for SVD-1 and SVD-2 subsamples together, hence the uncertainty is treated as completely correlated between them.

4. Proper decay time distribution

To check the reliability of efficiencies ϵ_{RS}^i and ratios $\epsilon_{\text{RS}}^i/\epsilon_{\text{WS}}^i$, and to estimate the effect of the imperfect fit to the proper decay time distribution, the values of ϵ_{RS}^i are compared to an alternative estimate from the fit to ΔM , $\epsilon_{\text{RS}}^{i,\Delta M} = N_{\text{RS}}^i/N_{\text{RS}}^{\text{tot}}$. This method accounts for the influence of the associated signal in the t_{xy} distribution. In a majority of the t_{xy} subintervals, ϵ_{RS}^i and $\epsilon_{\text{RS}}^{i,\Delta M}$ typically agree within $\pm 2\%$, the largest discrepancies being -9.8% and $+4.4\%$. For the integrated $1.6 < t_{xy} < 9.0$ interval, they agree within 0.8% – 1.4% for the four subsamples.

To estimate the effect of the discrepancies, the relative difference between ϵ_{RS}^i and $\epsilon_{\text{RS}}^{i,\Delta M}$ is assigned as the relative uncertainty on $\epsilon_{\text{RS}}^i/\epsilon_{\text{WS}}^i$. Hence we reduce the six efficiency ratios simultaneously by this uncertainty and repeat the R_M calculation; we then increase the ratios by this uncertainty, and again recalculate R_M . The difference between the resulting R_M value and the default fit is quoted as the systematic uncertainty from this source. It is very small and can be found on line 4 of Table VIII.

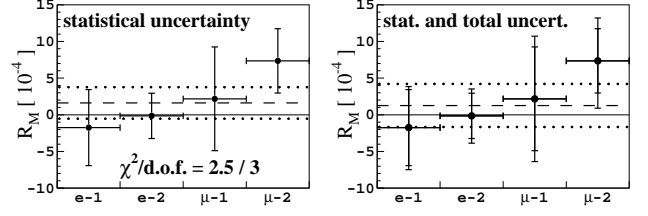


FIG. 16: Left: the R_M values of the four subsamples with the statistical uncertainty only, and the result of the fit to these four values (dashed line, $\chi^2/\text{d.o.f.} = 2.5/3$). Right: the four R_M values with the systematic uncertainty included and the combined result (dashed line), obtained as described in Sec. V A 7. The dotted lines represent the $\pm 1\sigma$ interval. The solid line corresponds to no mixing.

5. The amount of the associated signal

The systematic error due to the uncertainty in the associated signal fraction is estimated by varying the fraction and repeating the fitting procedure. Taking into account the uncertainties on the measured branching fractions [23] of the associated signal decay channels, we conservatively vary the amount of associated signal by $\pm 40\%$. We recalculate the R_M values and compare them to the default R_M value; we quote the differences as the systematic uncertainty from this source. From Table VIII (line 5) one can see that it is almost negligible.

6. The amount of the RS correlated background

From MC simulation studies we find that in the electron decay mode the largest contributions to the RS correlated background come from the following decays: $D^0 \rightarrow K^- \pi^+ \pi^0$ (33% of the RS correlated background), $D^0 \rightarrow K^- \pi^+ \pi^0 \pi^0$ (14%), $D^0 \rightarrow K^- e^+ \nu_e$ (13%). In total 60% of the RS correlated background comes from these three decays.

In the muon decay mode, the largest contributions to the RS correlated background come from $D^0 \rightarrow K^- \pi^+ \pi^0$ (43%), $D^0 \rightarrow K^- \pi^+ \pi^0 \pi^0$ (17%), $D^0 \rightarrow K^- \pi^+ \pi^- \pi^+$ (12%). In total 72% of the correlated background comes from these decays. We calculate the weighted average of the relative uncertainties of the branching fractions [23] for the stated decay modes. For the electron decay mode the averaged relative uncertainty is $\pm 4.3\%$ and for the muon decay mode $\pm 4.4\%$.

We repeat the RS fits, changing the amount of the total correlated background by the average uncertainties on the branching fractions. The differences between the obtained values of R_M and the default values are taken as the systematic uncertainty from this source. They can be found in Table VIII (line 6) and are negligible.

TABLE VIII: A summary of the systematic uncertainties on R_M , the total systematic uncertainty in each subsample and the combined (summed in quadrature) statistical and systematic uncertainty. Values are given in units of 10^{-4} .

source	e-1	e-2	$\mu-1$	$\mu-2$
1 fitting histo. statistics	± 1.54	± 0.91	± 2.64	± 1.81
2 WS correlated bkg.	± 0.37	$+0.39$ -0.38	$+2.98$ -2.89	$+3.05$ -2.97
3 WS uncorrelated bkg.	$+1.30$ -1.88	$+1.70$ -1.85	$+2.58$ -2.82	$+1.57$ -3.20
4 imperfect t_{xy}	± 0.05	± 0.02	± 0.01	$+0.25$ -0.33
5 associated signal	$+0.01$ -0.00	± 0.00	$+0.09$ -0.10	± 0.02
6 RS correlated bkg.	± 0.00	± 0.00	± 0.01	± 0.04
total systematic	$+2.05$ -2.46	$+1.97$ -2.10	$+4.75$ -4.83	$+3.89$ -4.74
statistical + systematic	$+5.58$ -5.74	$+3.66$ -3.73	$+8.53$ -8.57	$+5.86$ -6.45

7. Total systematic uncertainty and the final result

The final result of the measurement is obtained by averaging the results for the four subsamples, e-1, e-2, $\mu-1$ and $\mu-2$. As explained at the beginning of Sec. V, the results obtained by method 3) are used (quoted in Table VII) as “combined”.

The contributions to the systematic uncertainty are divided into three categories:

- The systematic uncertainty that is completely correlated between all four subsamples. This is the error due to the uncertainty of the WS correlated background fraction, Sec. V A 2.
- The systematic uncertainty that is completely correlated between the SVD-1 and SVD-2 subsamples and is uncorrelated between the electron and the muon decay mode. Such a contribution comes from the uncertainty of the ΔM shape of the uncorrelated WS background, Sec. V A 3.
- The systematic uncertainties that are uncorrelated between the four subsamples, or are very small. The main contribution comes from the uncertainty due to the finite statistics of the fitting distributions, Sec. V A 1. The uncertainties from all the remaining sources are also added.

To obtain the final result and its uncertainty, taking into account the systematic uncertainties, we adopt the following procedure:

- For each of the four subsamples, we add to the statistical uncertainty in quadrature all the uncertainties from category (c).
- We perform the χ^2 fit to the SVD-1 and SVD-2 R_M values in the electron and muon decay mode to obtain the averaged value for the electron and

muon decay mode, $(-0.56 \pm 2.76) \times 10^{-4}$ and $(5.89 \pm 4.02) \times 10^{-4}$, respectively. The quoted uncertainties include the statistical uncertainty and the uncertainties (c).

- To add the uncertainty (b) for the electron decay mode, we first simultaneously increase and then simultaneously decrease the results for the e-1 and e-2 subsamples by the uncertainty (b) and repeat step (2). The difference from the default result of step (2) is added in quadrature to the uncertainty obtained in step (2). The result for the electron decay mode, including the statistical uncertainty and systematic uncertainties (b) and (c), is $(-0.56^{+3.19}_{-3.33}) \times 10^{-4}$. We perform the same procedure also for the muon decay mode; the result is $(5.89^{+4.43}_{-5.07}) \times 10^{-4}$.
- We perform a χ^2 fit to the results for the electron and muon decay mode, obtained in step (3); the result is $(1.27 \pm 2.70) \times 10^{-4}$. The obtained mean value is the final result, but the uncertainty needs to be increased by the uncertainty (a).
- To account for the uncertainty (a) we first simultaneously increase and then simultaneously decrease the initial four R_M values by the uncertainties (a) and repeat the steps (1)–(4). The difference from the default result, $^{+1.13}_{-1.11} \times 10^{-4}$, is added in quadrature to the uncertainty previously obtained from step (4) to obtain the final uncertainty of the result.

The total uncertainty of the final result is $\pm 2.93 \times 10^{-4}$. We calculate the contribution of the systematic uncertainty as the difference between the total uncertainty and the statistical uncertainty (Eq. (22)), $2.93^2 - 2.16^2 = 1.98^2$. The final result is then

$$R_M = (1.3 \pm 2.2 \pm 2.0) \times 10^{-4}, \quad (23)$$

where the first uncertainty is statistical and the second systematic. As this value is close to the boundary of the physical region ($R_M \geq 0$) we use the Feldman-Cousins approach [26] to calculate upper limits:

$$R_M < 6.1 \times 10^{-4} \text{ at the 90\% confidence level,} \quad (24)$$

$$R_M < 7.0 \times 10^{-4} \text{ at the 95\% confidence level.} \quad (25)$$

With systematic uncertainties included, the final results for the electron and muon decay modes are:

$$R_M^e = (-0.6 \pm 2.7^{+1.8}_{-2.1}) \times 10^{-4}, \quad (26)$$

$$R_M^\mu = (5.9 \pm 3.7^{+3.9}_{-4.5}) \times 10^{-4}. \quad (27)$$

The R_M values for the four subsamples, including the systematic uncertainty, and the combined result are shown in the right plot in Fig. 16.

The increase in the sensitivity of the current result, compared to the one published in [7], is caused partially

by the larger statistical power of the sample, but also by the improvements in the measurement method. The statistical uncertainty of the present result in the electron sample (Eq. (26)) is about 22% smaller than one would expect by appropriately rescaling the uncertainty of the result [7] by the increase of the data set used. The improvement is mainly due to improved selection criteria, improved neutrino reconstruction and improvements in using the D^0 proper decay time measurement. The systematic error of the result in the electron sample is, however, larger than the one published in [7] as it is estimated more conservatively.

VI. SUMMARY

Using a data sample with an integrated luminosity of 492.2 fb^{-1} , collected by the Belle detector, we have searched for $D^0\text{--}\bar{D}^0$ mixing using semileptonic decays of the neutral charmed meson, $D^0 \rightarrow K^{(*)+}e^-\bar{\nu}_e$ and $D^0 \rightarrow K^{(*)+}\mu^-\bar{\nu}_\mu$. We select D^0 mesons produced via the decay $D^{*+} \rightarrow \pi_s^+ D^0$, and tag the flavor of the D meson at production by the charge of the accompanying slow pion. The measured mixing rate R_M is consistent with no mixing in both electron and muon decay modes. The combined result accounts for the partially correlated systematic error and yields $R_M = (1.3 \pm 2.2 \pm 2.0) \times 10^{-4}$. Since it is consistent with zero we set upper limits on the mixing rate of $R_M < 6.1 \times 10^{-4}$ at the 90% confidence level.

This result supersedes that published in Ref. [7] and represents the most stringent experimental limit on R_M obtained to date from semileptonic D^0 decays. Its accuracy is significantly better than that of the world average of previous measurements in semileptonic decays, $R_M = (1.7 \pm 3.9) \times 10^{-4}$ [4]. Although the sensitivity is not sufficient to observe a positive mixing signal, it is worth noting that in semileptonic decays no model uncertainties can influence the result. The reported value of R_M is in agreement with the world average values of $x = (0.87^{+0.37}_{-0.34})\%$ and $y = (0.66^{+0.21}_{-0.20})\%$ [4] and it will help in further constraining the D^0 mixing parameters in combination with the results of the measurements in other decay channels.

Appendix

In the proper decay time fit the following functions are used:

The Lorentz function, centered at 0:

$$L(t; b) = \frac{b}{\pi} \cdot \frac{1}{1 + (bt)^2} \quad (28)$$

The asymmetric Lorentz function, centered at 0:

$$L_a(t; b_l, b_r) = \begin{cases} \frac{b_l b_r}{(b_l + b_r)\pi} \cdot \frac{1}{1 + (b_l t)^2} & ; \quad t < 0 \\ \frac{b_l b_r}{(b_l + b_r)\pi} \cdot \frac{1}{1 + (b_r t)^2} & ; \quad t \geq 0 \end{cases} \quad (29)$$

Acknowledgments

We thank the KEKB group for the excellent operation of the accelerator, the KEK cryogenics group for the efficient operation of the solenoid, and the KEK computer group and the National Institute of Informatics for valuable computing and Super-SINET network support. We acknowledge support from the Ministry of Education, Culture, Sports, Science, and Technology of Japan and the Japan Society for the Promotion of Science; the Australian Research Council and the Australian Department of Education, Science and Training; the National Natural Science Foundation of China under contract No. 10575109 and 10775142; the Department of Science and Technology of India; the BK21 program of the Ministry of Education of Korea, the CHEP SRC program and Basic Research program (grant No. R01-2005-000-10089-0) of the Korea Science and Engineering Foundation, and the Pure Basic Research Group program of the Korea Research Foundation; the Polish State Committee for Scientific Research; the Ministry of Education and Science of the Russian Federation and the Russian Federal Agency for Atomic Energy; the Slovenian Research Agency; the Swiss National Science Foundation; the National Science Council and the Ministry of Education of Taiwan; and the U.S. Department of Energy.

-
- [1] M. Starič *et al.* (Belle Collaboration), Phys. Rev. Lett. **98** 211803 (2007).
 - [2] B. Aubert *et al.*, (BaBar Collaboration), Phys. Rev. Lett. **98**, 211802 (2007).
 - [3] See the talks by W. Lockman (http://chep.knu.ac.kr/lp07/htm/S4/S04_13a.pdf) and K. Tollefson (http://chep.knu.ac.kr/lp07/htm/S4/S04_14.pdf) at Lepton Photon 2007; W. M. Sun (for CLEO), arXiv:0712.0498v1.
 - [4] Heavy Flavor Averaging Group (Charm Decays subgroup), <http://www.slac.stanford.edu/xorg/hfag/charm/index.html>
 - [5] For a review see I.I. Bigi and N. Uraltsev, Nucl. Phys. B **592**, 92 (2001); S. Bianco, F.L. Fabbri, D. Benson and I. Bigi, Riv. Nuovo Cim. **26N7**, 1 (2003); A. Falk *et al.*, Phys. Rev D **69**, 114021 (2004).
 - [6] Z.-Z. Xing, Phys. Rev. D **55**, 196 (1997).
 - [7] U. Bitenc *et al.* (Belle Collaboration), Phys. Rev. D **72**, 071101 (2005).

- [8] B. Aubert *et al.* (BaBar Collaboration), Phys. Rev. D **70**, 091102 (2004); B. Aubert *et al.* (BaBar Collaboration), Phys. Rev. D **76**, 014018 (2007) .
- [9] C. Cawlfeld *et al.* (CLEO Collaboration), Phys. Rev. D **71**, 077101 (2005).
- [10] E. M. Aitala *et al.* (E791 Collaboration), Phys. Rev. Lett. **77**, 2384 (1996).
- [11] L. Zhang *et al.* (Belle Collaboration), Phys. Rev. Lett. **96**, 151801 (2006).
- [12] B. Aubert *et al.* (BaBar Collaboration), Phys. Rev. Lett. **97**, 221803 (2006).
- [13] B. Aubert *et al.* (BaBar Collaboration), hep-ex/0607090.
- [14] S. Kurokawa and E. Kikutani, Nucl. Instr. Meth. A **499**, 1 (2003), and other papers included in this volume.
- [15] A. Abashian *et al.* (Belle Collaboration), Nucl. Instr. Meth. A **479**, 117 (2002).
- [16] Z. Natkaniec *et al.* (Belle SVD-2 group), Nucl. Instr. Meth. A **560**, 1 (2006).
- [17] Events are simulated with the EvtGen generator, D.-J. Lange, Nucl. Instr. Methods Phys. Res., Sect. A **462**, 152 (2001); the detector response is simulated with GEANT, R. Brun et al., GEANT 3.21, CERN Report No. DD/EE/84-1, 1984.
- [18] The charge-conjugate modes are implied throughout the paper unless otherwise stated.
- [19] K. Abe *et al.* (Belle Collaboration), Phys. Rev. D **66** 032007 (2002).
- [20] G. C. Fox, S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978).
- [21] K. Hanagaki *et al.*, Nucl. Instr. Meth. A **485**, 490 (2002).
- [22] A. Abashian *et al.*, Nucl. Instr. Meth. A **491**, 69 (2002).
- [23] W.-M. Yao *et al.* (Particle Data Group), J. Phys. G **33**, 1 (2006).
- [24] In the following, P denotes the particle's 4-momentum while \vec{p} and p denote the 3-dimensional momentum and its magnitude, respectively.
- [25] The actual mass assignment of tracks included in P_{rest}^* has only a marginal effect on the neutrino reconstruction and thus on the ΔM resolution. If all charged particles are assigned a mass of the pion the ΔM resolution remains almost the same.
- [26] G.J. Feldman and R.D. Cousins, Phys. Rev. D **57**, 3873 (1998).